

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:34:53 ; Search time 47 Seconds

1 MNEPEQHRSMPKPKKNA (without alignments)

1983.752 Million cell updates/sec

Title: US-09-853-641-2

Perfect score: 1563

Sequence: 1 MNEPEQHRSMPKPKKNA.....LPAAVTWLILSVISSYPS 297

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	297	12	US-09-853-641-2
2	1495	95.6	285	9	US-09-738-626-5716
3	1495	95.6	285	12	US-10-627-476-100
4	1064	68.1	292	12	US-10-282-122A-53970
5	725.5	46.4	311	12	US-10-282-122A-62176
6	719	46.0	306	12	US-10-282-122A-62583
7	719	46.0	306	12	US-10-282-122A-64789
8	714	45.7	312	12	US-10-282-122A-63884
9	551	35.3	392	14	US-10-156-761-10160
10	310	19.8	271	12	US-10-282-122A-69792
11	308	19.7	285	12	US-10-282-122A-55991
12	306	19.6	79	11	US-09-824-408A-906
13	303	19.4	273	12	US-10-282-122A-49018
14	299	19.1	271	12	US-10-282-122A-67557
15	299	19.1	282	12	US-10-282-122A-77963

16	298	19.1	280	12	US-10-282-122A-77418	Sequence 77418, A
17	298	19.1	285	12	US-10-282-122A-75280	Sequence 75280, A
18	298	19.1	285	12	US-10-282-122A-76157	Sequence 76157, A
19	295	18.9	263	12	US-10-282-122A-59487	Sequence 59487, A
20	292.5	18.7	248	12	US-10-282-122A-73005	Sequence 73005, A
21	288.5	18.5	265	12	US-10-282-122A-65186	Sequence 65186, A
22	286.5	18.3	265	12	US-10-282-122A-65594	Sequence 65594, A
23	285	18.2	249	13	US-10-282-122A-56426	Sequence 56426, A
24	285	18.2	249	13	US-10-136-517-17	Sequence 17, Appl
25	283	18.1	265	12	US-10-282-122A-61125	Sequence 61125, A
26	282.5	18.1	287	12	US-10-282-122A-68659	Sequence 68659, A
27	280.5	17.9	289	12	US-10-282-122A-67464	Sequence 67464, A
28	277	17.7	288	12	US-10-282-122A-58351	Sequence 58351, A
29	276	17.7	273	12	US-10-282-122A-47625	Sequence 47625, A
30	276	17.7	273	12	US-10-282-122A-50187	Sequence 50187, A
31	257	16.4	274	12	US-10-282-122A-44770	Sequence 44770, A
32	249.5	16.0	260	12	US-10-282-122A-71462	Sequence 71462, A
33	249	15.9	241	12	US-10-282-122A-54574	Sequence 54574, A
34	248	15.9	241	12	US-10-282-122A-63044	Sequence 63044, A
35	247.5	15.8	260	12	US-10-282-122A-70940	Sequence 70940, A
36	247	15.8	293	15	US-10-369-493-16004	Sequence 16004, A
37	247	15.8	293	15	US-10-369-493-16356	Sequence 16356, A
38	247	15.8	309	15	US-10-369-493-15629	Sequence 15629, A
39	246.5	15.8	311	9	US-09-815-242-11873	Sequence 11873, A
40	246.5	15.8	311	12	US-10-282-122A-66387	Sequence 66387, A
41	244.5	15.6	264	12	US-10-282-122A-72396	Sequence 72396, A
42	243	15.5	285	12	US-10-282-122A-51478	Sequence 51478, A
43	239.5	15.3	266	12	US-10-282-122A-58551	Sequence 58551, A
44	237.5	15.2	260	12	US-10-282-122A-48529	Sequence 48529, A
45	237.5	15.2	296	15	US-10-369-493-19811	Sequence 19811, A

ALIGNMENTS

RESULT 1

US-09-853-641-2
; Sequence 2, Application US/09853641
; Publication No. US20040092710A1
; GENERAL INFORMATION:
; APPLICANT: NAMPOOTHIRI, Madhavan
; TITLE OF INVENTION: Nucleotide Sequences Coding for the CdsA Gene
; FILE REFERENCE: 032301 WD 1171
; CURRENT APPLICATION NUMBER: US/09/853,641
; CURRENT FILING DATE: 2001-08-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-853-641-2

Query Match 100.0%; Score 1563; DB 12; Length 297;
Best Local Similarity 100.0%; Pred. No. 3.6e-144;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MNEPEQHRSMPKPKKNAAGRLKAAIAVGLGLVLLGLVLSFWGWIYLVAGFMAA	60
Db	1	MNEPEQHRSMPKPKKNAAGRLKAAIAVGLGLVLLGLVLSFWGWIYLVAGFMAA	60
Qy	61	TWEVGSRLKEGGYHLPPIIMIGGQAIWLSPFPGTMGILASFVATVLVLMYFRIFYNGT	120
Db	61	TWEVGSRLKEGGYHLPPIIMIGGQAIWLSPFPGTMGILASFVATVLVLMYFRIFYNGT	120
Qy	121	EKEARNYLDTSVGIFVLTVILFSGFAAMLSIMQNNSTPGTYFILTFLMCLVIASDVGY	180
Db	121	EKEARNYLDTSVGIFVLTVILFSGFAAMLSIMQNNSTPGTYFILTFLMCLVIASDVGY	180
Qy	181	IAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHLLDHHWNGVILGALVYC	240
Db	181	IAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHLLDHHWNGVILGALVYC	240

QY 241 ATGLDVSQKRDIGIKDMSNLLPGHGLMDRLDGMPLPAAWTVLILSVISSYPS 297
 Db 241 ATGLDVSQKRDIGIKDMSNLLPGHGLMDRLDGMPLPAAWTVLILSVISSYPS 297

RESULT 2

US-09-738-626-5716
 ; Sequence 5716, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENO, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OKAZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patent In ver. 3.0
 ; SEQ ID NO 5716
 ; LENGTH: 285
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-5716

Query Match 95.6%; Score 1495; DB 9; Length 285;
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;
 Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 MPKPKNNAGRDLKAAIAGVIGLGLVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 72
 Db 1 MPKPKNNAGRDLKAAIAGVIGLGLVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 60

QY 73 YHLPPIIMIGQAIINLSWPFGTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 132
 Db 61 YHLPPIIMIGQAIINLSWPFGTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 120

QY 133 VGIFVLTWIPFGSPAAMLSLQNNISIPGTYFILTFLMCLVIASDVGGYIAGVFFGSHHPMA 192
 Db 121 VGIFVLTWIPFGSPAAMLSLQNNISIPGTYFILTFLMCLVIASDVGGYIAGVFFGSHHPMA 180

QY 193 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHWMMGVILGCALVVCATLGDLVESQFK 252
 Db 181 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHWMMGVILGCALVVCATLGDLVESQFK 240

QY 253 RDLGKMSNLLPGHGLMDRLDGMPLPAAWTVLILSVISSYPS 297
 Db 241 RDLGKMSNLLPGHGLMDRLDGMPLPAAWTVLILSVISSYPS 285

RESULT 3

US-10-627-476-100
 ; Sequence 100, Application US/10627476
 ; Publication No. US20040030116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompejus, Mark
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schoder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Habererhauer, Gregor

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 ; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
 ; FILE REFERENCE: BGI-125PCPN
 ; CURRENT APPLICATION NUMBER: US/10/627,476
 ; CURRENT FILING DATE: 2003-07-25
 ; PRIOR APPLICATION NUMBER: 09/602,787
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: USSN 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: DE 19931454.3
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931478.0
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931563.9
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19932122.1
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932180.9
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932124.8
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932125.6
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932128.0
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: DE 19932180.9
 ; PRIOR FILING DATE: 1999-07-09
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 678
 ; SEQ ID NO 100
 ; LENGTH: 285
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-627-476-100

Query Match 95.6%; Score 1495; DB 12; Length 285;
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;
 Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 MPKPKNNAGRDLKAAIAGVIGLGLVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 72
 Db 1 MPKPKNNAGRDLKAAIAGVIGLGLVLLGIVLSPWGHYILVAGFMAAATWEVGSRLKEGG 60

QY 73 YHLPPIIMIGQAIINLSWPFGTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 132
 Db 61 YHLPPIIMIGQAIINLSWPFGTWGLASVATVLMYFRIFYNGTEKEARNYLRDTS 120

QY 133 VGIFVLTWIPFGSPAAMLSLQNNISIPGTYFILTFLMCLVIASDVGGYIAGVFFGSHHPMA 192
 Db 121 VGIFVLTWIPFGSPAAMLSLQNNISIPGTYFILTFLMCLVIASDVGGYIAGVFFGSHHPMA 180

QY 193 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHWMMGVILGCALVVCATLGDLVESQFK 252
 Db 181 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHWMMGVILGCALVVCATLGDLVESQFK 240

QY 253 RDLGKMSNLLPGHGLMDRLDGMPLPAAWTVLILSVISSYPS 297
 Db 241 RDLGKMSNLLPGHGLMDRLDGMPLPAAWTVLILSVISSYPS 285

RESULT 4

US-10-282-122A-53970
 ; Sequence 53970, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Hasselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 53970

LENGTH: 292

TYPE: PRT

ORGANISM: Corynebacterium diptheriae

US-10-282-122A-53970

Query Match 68.1%; Score 1064; DB 12; Length 292;

Best Local Similarity 70.5%; Pred. No. 1.9e-95;

Matches 198; Conservative 33; Mismatches 50; Indels 0; Gaps 0;

13 MKPKNNAGRDILKAAIYAGIGLVLGIVLSPGWLIVAGFMAAATWVGSRLEGG 72

11 LPKKNAGRLKAAISVIGLGLVLAIFVIFGWTPLVAIAIATWVEERLLEAG 70

73 YHLPPLIMIGQAIIMLSWPFMTGILASVATVLMYRIFNYNGTEKARNYLRDTS 132

71 YLLQSRWMLIGQVNLMSWPFPGKGLVAGFVGVVLTATMFCRLPHGSRMPPKNYLRDTA 130

133 VGIFVLTWIPFGSFAAMLSLMQNSISCTYFILTFLMCLVTSADVGGYIAGVFGSHPMA 192

131 VAIFVLTWIPFGSFAAMLSLFTETAPGKYPIVTFMLCVIASDVGGYIAGVFGSHPMA 190

193 PLVSPKSWEGFAGSIVLGSVTGALSVEHFLDHHWMMGVILGCALVVCATLGLDLVESQFK 252

191 PAVSPKSWEGFAGSIVLGSVTGALSVEHFLDHHWMMGVILGCALVVCATLGLDLVESQFK 250

253 RDLGKQNSNLLPHGGMLMDRLDGMPLPAAMTWLILSVISS 293

251 RELGKQNSAAILPHGGMLMDRLDGMPLPAAMTWLILSVISS 291

RESULT 5

US-10-282-122A-62176

Sequence 62176, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 62176

LENGTH: 311

TYPE: PRT

ORGANISM: Mycobacterium avium

US-10-282-122A-62176

Query Match 46.4%; Score 725.5; DB 12; Length 311;

Best Local Similarity 47.5%; Pred. No. 2.4e-62;

Matches 145; Conservative 47; Mismatches 94; Indels 19; Gaps 4;

2 NEPE-----QHRSRMPPKPKNNAGRDILKAAIYAGIGLVLGIVLSPGWYILVA 54

10 DEPEHAVENTTEGAAGQRAKKTSRAGDRDLRAAIVAGAGIATVLTVPAPFWPIVA 69

55 GFMAAATWVGSRLEKGGYHLPIMIGQAIIMLSWPFMTGILASVATVLMYR 114

70 MALIVASHEVVRLEAGYVIVIPILLAGGQLTWLTWPFHAAGALAGFGTVVACLFWR 129

115 IFYNGTEKE-----ARNYLRDTSVGIFVLTWIPFGSFAAMLSLMQNSISCTYFILTFLMCLVTSADVGGYIAGVFGSHPMA 166

130 LFQMDNRKRPEPPAGSPSANYLRDASATVFLACWVPLFASFAALLVYPADGA--GRVFCL 187

167 TFMCLVTSADVGGYIAGVFGSHPMAPLVSPKSWEGFAGSIVLGSVTGALSVEHFLDHH 226

188 --NITVTSADVGGYIAGVFGSHPMAPLVSPKSWEGFAGSIVLGSVTGALSVEHFLDHH 245

227 WMMGVILGCALVVCATLGLDLVESQFKRDLGKQNSNLLPHGGMLMDRLDGMPLPAAMTWL 286

246 PWVGALLGVVLTCTGLDLVESQVKRDLGKQNSNLLPHGGMLMDRLDGMPLPAAMTWL 305

287 ILSVI 291

306 VLTIV 310

RESULT 6

US-10-282-122A-62583

Sequence 62583, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

```

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62583
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62583

Query Match 46.0%; Score 719; DB 12; Length 306;
Best Local Similarity 47.7%; Pred. No. 1e-61;
Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;

QY 2 NEPEQHRSRM--PKPKNNAGRDLEKAAIAGVIGLGLVLLGLVLSPGMGWYILVAGFMAA 59
DB 10 NPAEQPARAKQOPATETSRAGRDLEKAAIAGVIGLGLVLLGLVLSPGMGWYILVAGFMAA 69
QY 60 ATWEVGSRLKEGGYHLPIMIGQAAIWLSPFGTNGILASFVATVLMYPRIFY-- 117
DB 70 ATHEVVRRLREAGYILPVIPLLIQGAQAVLTPFGAVGALAGFGGMVVCMIWRLEFMD 129
QY 118 -----NGTEKEARNYLDTSVGIPLTWIPLFGSFAAMLQNNNSIPGYFITLTMLC 171
DB 130 SVTRPTTGAPSPGNYSVTSATVLAIVVPLFCFGLAVLPEN----GSGWVFCMMIA 185
QY 172 VIASDVGGYIAGVFFGSHPMAPLSPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 231
DB 186 VIASDVGGYAVGVLFQKHPVPTISPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 245
QY 232 ILGCALVVCATIGDILVESQFKRDIGIKDMSNLLPGHGLMDRLDGLMPLPAAMVTWILSVI 291
DB 246 LLGVLFVLTALGDILVESQVKRDIGIKDMSNLLPGHGLMDRLDGLMPLPAAMVTWILSVI 305

RESULT 7
US-10-282-122A-64789
; Sequence 64789, Application US/10282122A
; Publication No. US20040029129A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 64789
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64789

Query Match 46.0%; Score 719; DB 12; Length 306;
Best Local Similarity 47.7%; Pred. No. 1e-61;
Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;

QY 2 NEPEQHRSRM--PKPKNNAGRDLEKAAIAGVIGLGLVLLGLVLSPGMGWYILVAGFMAA 59
DB 10 NPAEQPARAKQOPATETSRAGRDLEKAAIAGVIGLGLVLLGLVLSPGMGWYILVAGFMAA 69
QY 60 ATWEVGSRLKEGGYHLPIMIGQAAIWLSPFGTNGILASFVATVLMYPRIFY-- 117
DB 70 ATHEVVRRLREAGYILPVIPLLIQGAQAVLTPFGAVGALAGFGGMVVCMIWRLEFMD 129
QY 118 -----NGTEKEARNYLDTSVGIPLTWIPLFGSFAAMLQNNNSIPGYFITLTMLC 171
DB 130 SVTRPTTGAPSPGNYSVTSATVLAIVVPLFCFGLAVLPEN----GSGWVFCMMIA 185
QY 172 VIASDVGGYIAGVFFGSHPMAPLSPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 231
DB 186 VIASDVGGYAVGVLFQKHPVPTISPKKSWEFGAGSIVLGSVTGALS VHFLLDHHMMGV 245
QY 232 ILGCALVVCATIGDILVESQFKRDIGIKDMSNLLPGHGLMDRLDGLMPLPAAMVTWILSVI 291
DB 246 LLGVLFVLTALGDILVESQVKRDIGIKDMSNLLPGHGLMDRLDGLMPLPAAMVTWILSVI 305

RESULT 8
US-10-282-122A-63884

```



```

RESULT 10
US-10-282-122A-69792
/ Sequence 69792, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeek, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Bases
/ FILE REFERENCE: ELIPIRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20

```

```

; PRIOR APPLICATION NUMBER: 60/191,078
; CURRENT FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69792
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
; US-10-282-122A-69792

```

```

Query Match      19.8%; Score 310; DB 12; Length 271;
Best Local Similarity 33.3%; Pred. No. 7,7e-22;
Matches 99; Conservative 44; Mismatches 84; Indels 70; Gaps 17;

QY 22 RLKAAIANGIGL-GVVLVLG-----IVLSPWGWYILVAGFMAAATWVG----- 65
DB 5 RIITALLIPALCGFFLLTGTMYFALF-GVVVVLGAWEWARL-AGF-AAQSMRVGYAALV 62
QY 66 SRLKEGGVHLP-LPMIIGGOAIIWLSWPFGTWGIILASPVATVLVLMYFRIFYNGTEKE 123
DB 63 AVLLFPWYLLPGLFWLVA--AVIWS-----VATFLVLY-----PD 99
QY 124 ARNYLEDTS---VGIFVLTWTFPGSPAAMLNMONNSIP-GTYFIITFMCLVIADSVG 178
DB 100 SSSHWASAACKLVIGLLIL-----LPAMQGLVLKQWPLGNWLLLSVAVLYWAADIG 151
QY 179 GYIAGVFFGSHPMAPLVSPPKSNWEGFAGSIVLG-SVTGALSVMHFLDDHW-----WVGVL 233
DB 152 AVFSGKAFGRKRLAPKVSPPKSNWEGYGLVLSLGITAAVGV-----RDWTVVGFIAALL 207
QY 234 GCALVV-CATLGDVLESQPKDLGKIDMSNLLPGRHGLMDRLDGLMPLA-----AMVTW 285
DB 208 GAIVVIFSVIGDLTESMFKRGSKGVKDSNLLPGRHGLVDRIDSLTAIPVFAVLLW 264

```

```

RESULT 11
US-10-282-122A-55991
; Sequence 55991, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A

```

```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55991
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
; US-10-282-122A-55991

```

```

Query Match      19.7%; Score 308; DB 12; Length 285;
Best Local Similarity 30.7%; Pred. No. 1.3e-21;
Matches 89; Conservative 47; Mismatches 92; Indels 62; Gaps 13;

QY 30 VGTGLGVLLGLVLSPWG-----WYLVAGFMAAATWVGSRLEGGVHLP 77
DB 26 VGFAIVTVLVVCLAAWEGQFSGFTSRQVRLVLCGFLAI--MLFTLPE--YHDDI 80
QY 78 PIMIIGQAIWLSWPFGTWGIILASPVATVLVLMYFRIFYNGTEKEARN--YLRDTSVGI 135
DB 81 HQPELVAGS--LWIS-----LAWWVAALLV---LFPGSASLWRNSKVLRL--LI 122
QY 136 F-VLTWIPFGSPAAMLN-MQNNSTPGYFYFILTMCVVIADSVGYIAGVFGSHPMAP 193
DB 123 FGLLTIIPTFFWGMVALRAWHYDENHYSGAIIWLLVYVILVWGADSGAYMFGKLFGRHKLAP 182
QY 194 LVSPKSNWEGFAGSIVLG-SVTGALSVMHFLDDHWVG-----ILGAL--VUCA 241
DB 183 KVSFGKTFWQGFGLFTAAIIS-----WGVGVWANLEVAAPSILIVCSIPAALAS 231
QY 242 TLGDVLESQPKDLGKIDMSNLLPGRHGLMDRLDGLMPLAAMVTWILSVI 291
DB 232 VLGDLTESMFKREAGIKDSCHLIPGRHGLVDRIDSLTAAPVFAVLLLV 281

```

```

RESULT 12
US-09-864-408A-906
; Sequence 906, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encoc
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 906
; LENGTH: 79

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:32:48 ; Search time 16 Seconds
(without alignments)
1785.555 Million call updates/sec

Title: US-09-853-641-2

Perfect score: 1563

Sequence: 1 MNEPEQHRSRMMPKPKNA.....LPAAMVTWLILSVSSYPS 297

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:**

1: Pir1:**

2: Pir2:**

3: Pir3:**

4: Pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	719	46.0	306	D70924	probable cdaA prot
2	714	45.7	312	G87107	probable phosphati
3	539	34.5	391	T35470	probable integral
4	303	19.4	271	JC4832	phosphatidate cyti
5	299	19.1	282	A10128	phosphatidate cyti
6	298	19.1	271	F83188	phosphatidate cyti
7	298	19.1	280	E82099	phosphatidate cyti
8	298	19.1	285	A30529	phosphatidate cyti
9	286.5	18.3	265	A82000	phosphatidate cyti
10	285	18.2	249	1 SYECDG	phosphatidate cyti
11	285	18.2	249	A39651	phosphatidate cyti
12	285	18.2	249	A35502	CDP-diglyceride sy
13	280.5	17.9	265	B11229	CDP-diglyceride sy
14	277	17.7	288	G54102	phosphatidate cyti
15	265.5	17.0	294	A22290	phosphatidate cyti
16	259.5	16.6	284	B75386	phosphatidate cyti
17	253	16.2	399	T47873	phosphatidate cyti
18	252.5	16.2	430	A84887	probable phosphati
19	251.5	16.1	259	A70408	phosphatidate cyti
20	249	15.9	241	H81278	phosphatidate cyti
21	248.5	15.9	278	A83355	phosphatidate cyti
22	246.5	15.8	311	A83329	probable phosphati
23	245.5	15.7	310	H82138	probable phosphati
24	242.5	15.5	277	AC2746	phosphatidate cyti
25	242.5	15.5	277	B37527	hypothetical prote
26	240.5	15.4	269	G69597	phosphatidate cyti
27	239.5	15.3	266	G54546	CDP-diglyceride sy
28	238.5	15.3	312	G70114	phosphatidate cyti
29	236.5	15.1	260	H89899	phosphatidate cyti

30	236.5	15.1	298	2	C90880	probable phosphati
31	236.5	15.1	298	2	F85738	probable phosphati
32	235.5	15.1	284	2	F83952	phosphatidate cyti
33	234.5	15.0	284	2	A82729	phosphatidate cyti
34	231.5	14.8	266	2	F71961	cdp-diacylglycerol
35	231	14.8	275	2	A87487	phosphatidate cyti
36	230.5	14.7	262	2	AH1601	phosphatidate cyti
37	230.5	14.7	298	2	D64892	phosphatidate cyti
38	228.5	14.6	262	2	AD1239	phosphatidate cyti
39	225.5	14.4	267	2	A97902	phosphatidate cyti
40	224	14.3	293	2	S77254	phosphatidate cyti
41	223	14.3	230	2	E97773	hypothetical prote
42	219.5	14.0	267	2	G95030	phosphatidate cyti
43	216.5	13.9	228	2	G71700	phosphatidate cyti
44	206	13.2	378	2	E82882	CDP-diglyceride sy
45	204	13.1	287	2	D71304	probable phosphati

ALIGNMENTS

RESULT 1

D70924

Probable cdaA protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C/Accession: D70924

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: D70924

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-306 <COL>

A/Cross-references: GB:Z74024; GB:AL123456; NID:G3250700; PIDN:CAA98357.1; PID:G1403401

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: cdaA

C/Superfamily: phosphatidate cytidyltransferase

Query Match 46.0%; Score 719; DB 2; Length 306;

Best Local Similarity 47.7%; Pred. No. 2.1e-54;

Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;

QY	2	NEPEQHRSRM--PKPKNAGRDLCAAATAVGIGLVGLVLSVPGWGYILVAGFMAA	59
DB	10	NPAEQPARAKQOPATETSRGRDLRAIVVGLSIGLVIAVLVFRVVAIVAVATLV	69
QY	60	ATWEVGRLEKEGYHLPPLPIMIGGOAILWSPFGTGMIGILASFVATVLVLMYFRIFY--	117
DB	70	ATHEVVRRLREAGYLIPVPLLLIGGQAAVLTWPFCAVAGLAFGGVMVVCMIWRLFMQD	129
QY	118	-----NGTEKEARNYLRDTSVGIFVLVTPLEGSFAAMLISLMONNSIDGTFFILTFMLC	171
DB	130	SVTRPTTGAPSPGNYLSVTSATVFLAVWVPLFCSEFGLVVPEN---GSGWVFCNMIA	185
QY	172	VIASDVGGYIAGVFFGSHPMAPLVSPKKSWEQFAGSIVLGSVTGALSVHFLDHHMMGV	231
DB	186	VIASDVGGYAVGVLFQKHPMVFTISPKKSWEQFAGSLVCGITATITATFLVGTWIGA	245
QY	232	ILGCALVCAVATGDIVESQFKRDLGDKMNSLLPHGGMLMDRLDGMPLPAAMVTWLILSVI	291
DB	246	LLGVLFVLTALGDLVSVQKRDGDKMGRLLPHGGMLMDRLDGLTLPSSAIVLITLL	305

RESULT 2

G87107

probable phosphatidate cytidyltransferase [imported] - Mycobacterium leprae
C/Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-May-2002
C;Accession: G87107
R;Cole, S.T.; Biglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hc
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; PMID:21128732; PMID:11234002
A;Accession: G87107
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <STO>
A;Cross-references: GB:AL450380; NID:gl3093385; PIDN:CAC30540.1; GSPDB:GN00147
C;Genetics:
C;Gene: ML1589
C;Superfamily: phosphatidate cytidyltransferase
Query Match 45.7%; Score 714; DB 2; Length 312;
Best Local Similarity 49.5%; Pred. No. 5.8e-54;
Matches 146; Conservative 49; Mismatches 84; Indels 16; Gaps 5;
QY 9 RSMRW-PKPKNAGRDLCALAAVAGIGLVLLGIVLSPWGMVILVAGFMAAAATWEVGRS 67
DB 21 RAMROSTKNTPRAGNLPAIAVAGLSIGGLVATLVAFAPRIWVVLCAIAIFVAGHEVVR 80
QY 68 LKEGGYHLPPLPMIIGGQAIWLSWPPFGTMGLASFAVATVLMVFRIFY--NGTEKEAR 125
DB 81 LREAGYVPAIPALIGGQFTWLTWTPYTVGALAGFGATVVCWIMRWMDNSKQHSR 140
QY 126 -----NYLRDTSVGIFFVLTWPLFGSPAAMLSLMQNNISIPGTFFILTMCLVIA 176
DB 141 BALAGPPVSNVLRDASATVFLAAWVPLFASPAALLVPEKGA--GRVFCL--MIAVVASD 196
QY 177 VGGYTAGVFFGSHMPALVSPKKSNEGFAGSVLGSVTGALSVHFLDHHWGMVILGCA 236
DB 197 VGGYTAGVFLGKHPLVPRISPNKSEGFAGSVLGSVTGATLTATLTATFLAGKTPWVGALLSFV 256
QY 237 LVVCATIGDLVESQKRDIGIKMNSNLLPGHGLMDRLDGLMPLPAAMVTLILSVI 291
DB 257 LVLTCTGLDLVESQKRDIGIKMNSNLLPGHGLMDRLDGLMPLPAAMVTLILSVI 311
RESULT 3
T35470
probable integral membrane protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C;Accession: T35470
R;Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1998
A;Reference number: Z21578
A;Accession: T35470
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-391 <MUR>
A;Cross-references: EMBL:AL031035; PIDN:CAA19924.1; GSPDB:GN00070; SCODEB:SC6A9.39c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCODEB:SC6A9.39c
Query Match 34.5%; Score 539; DB 2; Length 391;
Best Local Similarity 38.8%; Pred. No. 8.2e-39;
Matches 114; Conservative 51; Mismatches 117; Indels 12; Gaps 5;
QY 4 PQOHRSMPKP-KNNAGRDLCALAAVAGIGLVLLGIVLSPWGMVILVAGFMAAAATW 62
DB 107 PDAQPQSQPPPKKSAGRDLCALAAVAGIGLVLLGIVLSPWGMVILVAGFMAAAATW 166
QY 63 EYGSRIKE-GGYHLPPLPMIIGGQAIWLSWPPFGTMGLASFAVATVLMVFRIFYNGTE 121
DB 167 ELTKLEERKGIKAPLVPLAIGGAAMVAGYARGAAGAWAMALTALAVLVRM-----T 221

QY 122 KEARNYLRTSVGIFVLTWPLFGSPAAMLSLMQNNISIPGTFFILTMCLVIA SDVGGYI 181
DB 222 EPPEGYLKDVITAGLFAAFVFPFLATFVAMVLAADD-----GAWRVLVFLILTVSTGTAYA 277
QY 182 AGVFFGSHMPALVSPKKSNEGFAGSVLGSVTGALSVHFLDHHWGMVILGCAVVC 240
DB 278 VGMRFGRKKLAPRISPGKTREGLLGAIAFAMVAGALCMQFLIDDCGAWWQGLLLGLWAVS 337
QY 241 ATLGLDVLVESQKRDIGIKMNSNLLPGHGLMDRLDGLMPLPAAMVTLILSVISSS 294
DB 338 ATLGLDGLGESMIKRDIGIKMNSNLLPGHGLMDRLDGLMPLPAAMVTLILSVIFVGS 391
RESULT 4
JC4832
phosphatidate cytidyltransferase (EC 2.7.7.41) - Pseudomonas aeruginosa
N;Alternate names: CDP-diglyceride synthetase
C;Species: Pseudomonas aeruginosa
C;Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 20-Jun-2000
C;Accession: JC4832
R;Taguchi, K.; Fukutomi, H.; Kuroda, A.; Kato, J.; Ohtake, H.
Gene 172, 165-166, 1996
A;Title: Cloning of the Pseudomonas aeruginosa gene encoding CDP-diglyceride synthetase.
A;Reference number: JC4832; PMID:96257274; PMID:8654980
A;Accession: JC4832
A;Molecule type: DNA
A;Residues: 1-271 <TAG>
A;Cross-references: DDBJ:D50811; NID:gl262331; PIDN:BA049437.1; PID:gl262332
C;Comment: This enzyme catalyses the condensation of CTP and phosphatidic acid to form CI
C;Genetics:
C;Superfamily: phosphatidate cytidyltransferase
C;Keywords: nucleotidyltransferase
Query Match 19.4%; Score 303; DB 2; Length 271;
Best Local Similarity 33.3%; Pred. No. 1.1e-18;
Matches 99; Conservative 39; Mismatches 93; Indels 64; Gaps 15;
QY 22 RDLKAAIAVAGIGLVLL-----GIVLS--PWGVIILVAGF-----MAAAATWE 63
DB 5 RIITALLVPLALGGFFLLEGAFALFICAVVSLGAWENARL-AGVEQOQGRVAVAAATVA 63
QY 64 VGSRLKEGGYHLPPLPMIIGGQAIWLSWPPFGTMGLASFAVATVLMVFRIFYNGTEKE 123
DB 64 V---LMVALYHLP---QLAGAVLLGLAVW-----WTLATVLTLY-----PE 99
QY 124 ARNYL--RDTSVGIFVLTWPLFGSPAAMLSLMQNNISIPGTFFILTMCLVIA SDVGGYI 181
DB 100 SVGYWGRWRRLGMLLILLPAWQGLVLLKQVAAANGL-----IIAVVVLWGADIGAYF 154
QY 182 AGVFFGSHMPALVSPKKSNEGFAGSVLGSVTGALSVHFLDHHWGMVILGCA 236
DB 155 SGKAFGRKKLAPRISPGKTREGLLGAIAFAMVAGALCMQFLIDDCGAWWQGLLLGLWAVS 210
QY 237 LVVCATL-GDLVESQKRDIGIKMNSNLLPGHGLMDRLDGLMPLA-----AMVTW 285
DB 211 LVVTVSIVGDLTIESMPKRSKIKSSNLLPGHGLVLDRLDLSLTAAIPVFAALLW 264
RESULT 5
AI0128
phosphatidate cytidyltransferase (EC 2.7.7.41) [imported] - Yersinia pestis (strain COI
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI0128
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;
delo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; PMID:21470413; PMID:11586360
A;Accession: AI0128
A;Status: preliminary

[illegible]

Db 7 WQGLSGFTTRSRQVWLAVCGLLLLALMLFL---LPEYHRNHQPLUVEITS-----LWASLW 59

Qy 92 WPFGTMGILASFVATVVLVLMYPRIFYNGTEKEARNYLRDTSVGIF-VLTWIPLF-GSPAA 149

Db 60 W-----IVALLVL-----FYPGSAALWRN--SKTLRLIFGULTIVPFPGWGLAL 102

Qy 150 MSLMLONNSIPGTYFILTFLMCLVLIADVGGYIAGVFFGSHPMAPLVSPKKSWEFGAGSIV 209

Db 103 RAWHYDENHYSAGIALLVYMLVLWGLADSGAYFGKLFQGHKLAPKVSFGKTWQGFIGGLA 162

Qy 210 LGSVTCALSVHFLLDHHMMWGV-----ILGCVLV--VCATLGLDVLVESQFKRDLGI 257

Db 163 TAAVIS-----NGYGMKANLDVAPVTLTLCISIVAALASVLGDLTESMFKREAGI 211

Qy 258 KMSNLLPHGGMLDRDLGMLPAAMVTWLLISVI 291

Db 212 KDSGHLIPHGGILDRIDSLTAAVVFACLLLV 245

RESULT 13

B81229
phosphatidate cytidylyltransferase NMB0185 [imported] - Neisseria meningitidis (strain MC58)
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: B81229
R/Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Xie, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Verma, A.; Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A/Reference number: A81000; MUID:20175755; PMID:10710307
A/Accession: B81229
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-285 <RET>
A/Cross-references: GB:AE002375; GB:AE002098; NID:g7225394; PID:AAF40642.1; PID:g7225404.1
A/Experimental source: serogroup B, strain MC58
C/Genetics:
A/Gene: NMB0185
C/Superfamily: phosphatidate cytidylyltransferase

Query Match 17.9%; Score 280.5; DB 2; Length 265;
Best Local Similarity 28.9%; Pred. No. 9e-17;
Matches 83; Conservative 50; Mismatches 113; Indels 41; Gaps 9;

Qy 24 LKAAATAGVIGLGVVLIGLIVLSP-WGVYILVAGFMFAATWE---VGSRLKEGGVHLPLPI 79

Db 2 LQRVITAWLLPLMLGLMFYAPQWLNAAFGLTIALIWLWYARMGGICKIKTHYLAAT 61

Qy 80 MITGOA-----IWLSPFGTNGILASFVATVVLVLMYPRIFYNGTEKEARNYLRD 130

Db 62 LVFGVVAVAGGWMLPNLVYV-----VLAFLVAVMPLWLRFKWRLNGGQ----- 106

Qy 131 TSVGIFVLTWLPFGSFAAMLISM--QNNSTPGTYFILTFLMCLVLIADVGGYIAGVPEGS 188

Db 107 ----VYAGWLVNWFVFWALSLRPHDPDLP-----LLAVMGLVWVADICAYFGSKAFGK 158

Qy 189 HPMAPLVSPKKSWEFGAGSIVLGSV--TCALSVHFL-LDHEWMMGVILGCVLWVATLGD 245

Db 159 HKIAPAIISFGKSWEAGVAVYVYAVRSAGNLAFDTGWFDTVLIGLVLTVVSVCGD 218

Qy 246 LVESQFKRDLGIKDSNLLPHGGMLDRDLGMLPAAMVTWLLISVIS 292

Db 219 LLESWLKRAAGIKDSKLLPHGGHGVDFRDTDSLIAVISVYAAWMSVLN 265

RESULT 14

G64102
phosphatidate cytidylyltransferase homolog - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 21-Jan-2000
C/Accession: G64102

61 QY -----TWEVGSRLKEGGYHLPIMIGGQAIWLSPFOTMGILASFVATVVLV 110
 58 Db TLFLSQVLLVTCITLDSNLVDA-----VMPAGTFICFYLLFQPKMATIADVSASIMGL 110
 111 QY MYFRIFYNGTEKEARNYLRDTSVGIFVLTWIPLFGSFAAMLS--LMQNN--SIPQTY--F 164
 111 Db FVGYL-----PSYVWLRSIGSTSISNPLGGYWGSGTDLLEERNFASLFGQFKIT 163
 165 QY ILTFMLCVIASDVGGYIAGVFFGSHPMAPLVSPKKSWEG---FAGSIVLG-SVTGALS-VH- 220
 164 Db VLTFF-LCIWAADIGAYTIGKFFGKTRLS-D-ISPKKTVEGAVEGISASLAVAGAFYLHL 221
 221 QY ---FLDHHWMMGVILGCALVVCATLGDIVESQFKDGLGKMSNLLPGHGLMDRLDGM 278
 222 Db POFLL-----TGTHLGLLIGLASLLGLTESMLKRDAGVKDGGQLIPGHGGLDKTDSYI 276
 279 QY PAAMVTWLILSVI 291
 277 Db FTAPLVYVYFVTLI 289

Search completed: August 11, 2004, 13:36:26
 Job time : 16 secs

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
 ; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 436-512, 1995
 A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A;Reference number: A64000; MUID:95350630; PMID:7542800
 A;Accession: G64102
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-288 <1000>
 A;Cross-references: GB:U32773; GB:L42023; NID:G1573932; PIDN:AA22577.1; PID:G1573940;
 C;Superfamily: phosphatidate cytidylyltransferase

Query Match 17.7%; Score 277; DB 2; Length 288;
 Best Local Similarity 28.9%; Pred. No. 2e-16;
 Matches 87; Conservative 32; Mismatches 76; Indels 106; Gaps 13;
 32 IGLGVLLVGLVILSPWGH-----YILVAGFMAA--ATWEVGSRLKEGGY----- 73
 30 LALGAVAILGI-----WEWTFARLQPLIRFPVTFLLGVFFLW-----LYTEGNYLDAGR 81
 74 ---HLPLPIMIGQAIIW-----LSWP-----FQTMGILASFV 104
 82 VFEHLQLLLI-----NAVSWGLALLLVISYPSKAKFWSKNPLQLLFAFST---LIPFV 134
 105 ATVLVIMYFRIFYNGTEKEARNYLRDTSVGIFVLTWIPLFGSFAAMLSLMQNNISIPQTYF 164
 135 AGVLR-----RLEHYTHDPYHGLFLLLV-----F 160
 165 ILTFMLCVIASDVGGYIAGVFFGSHPMAPLVSPKKSWEGAGSIVLGSVTCALS-VHF--- 221
 161 IL-----VWAADGAYFGRAPGKRLAPKVPKSGKVEGVTGGITLVLAIFIHFSNN 215
 222 ---LDPHWWMMGVILGCALVVCATLGDIVESQFKDGLGKMSNLLPGHGLMDRLDGM 279
 216 TLVGRNITGFIILSVATVATVLSVLGDLTFESMFKRESGVKDSQLIPGHGGLDRIDSLTA 275
 280 A 280
 276 A 276

RESULT 15
 AD2290
 phosphatidate cytidylyltransferase [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AD2290
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 203-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AD2290
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-294 <NUP>
 A;Cross-references: GB:BA000019; PIDN:BA075574.1; PID:G17133009; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: all3875
 C;Superfamily: phosphatidate cytidylyltransferase

Query Match 17.0%; Score 265.5; DB 2; Length 294;
 Best Local Similarity 29.4%; Pred. No. 2e-15;
 Matches 92; Conservative 49; Mismatches 101; Indels 71; Gaps 15;
 22 RDLKAAIAGVIGLGVLLGLVILSPWGYILV-----AGFMAA-- 60
 5 RIISGIVAILAL-VAVLLG-----GWYFTIMLAIIVVLGQBYFDLVRTRGILPAKT 57

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		ID	Description
		Match	Length		
1	719	46.0	306	1	CDSA_VYCTU
2	714	45.7	312	1	CDSA_VYCLE
3	298	19.1	271	1	CDSA_PSEAE
4	285	18.2	249	1	CDSA_ECOLI
5	277	17.7	288	1	CDSA_HAEIN
6	251.5	16.1	259	1	CDSA_BACSU
7	244	15.6	270	1	CDSA_AQUAE
8	243.5	15.6	270	1	CDSA_BRUAB
9	240.5	15.4	269	1	CDSA_BRUSU
10	239.5	15.3	266	1	CDSA_HELPY
11	231.5	14.8	266	1	CDSA_HELPU
12	230.5	14.7	298	1	YNBB_ECOLI
13	224	14.3	293	1	CDSA_SYNT3
14	216.5	13.9	228	1	CDSA_RICPR
15	202	12.9	270	1	CDSA_THEMA
16	201.5	12.9	447	1	CDSA_THROW
17	188.5	12.7	445	1	CDS2_HUMAN
18	197.5	12.6	424	1	CDS1_SOLTU
19	197	12.6	421	1	CDS1_ARATH
20	195.5	12.5	461	1	CDS1_HUMAN
21	194.5	12.4	461	1	CDS1_RAT
22	193.5	12.4	308	1	CDSA_CHLPN
23	188.5	12.1	305	1	CDSA_CHLFR
24	180.5	11.5	305	1	CDSA_CHLMD
25	177	11.3	455	1	CDS1_CABEL
26	169.5	10.8	395	1	CDSA_MYCPN
27	161	10.3	305	1	CDSA_MYCGE
28	142	9.1	457	1	CDS1_YEAST
29	119	7.6	136	1	CDS1_MOUSE
30	115.5	7.4	326	1	MRAY_STREN
31	115.5	7.4	326	1	MRAY_STRRG
32	113	7.2	417	1	CLOC_SALTI
33	112.5	7.2	329	1	MRAY_LACIA


```

RESULT 3
CDS: PSEAE
ID CDS: PSEAE STANDARD; PRT: 271 AA.
AC Q59640;
AT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 42, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
DE synthase) (CDS) (CTP-phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
GN CDS OR CDS OR PA3651.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN Gene 172:165-166(1996).
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=96257274; PubMed=8654980;
RA Taguchi K., Fukutomi H., Kuroda A., Kato J., Ohtake H.;
RT "Cloning of the Pseudomonas aeruginosa gene encoding CDP-diglyceride
RT synthetase.";
RL Gene 172:165-166(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.;
RA Hickey M.J., Brinkman F.S.L., Hufragle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
[1]
CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
CC diacylglycerol.
CC -!- PATHWAY: Phospholipid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: Belongs to the CDS family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D50811; BRA09437.1; -.
CC EMBL; AB004785; NAG07039.1; -.
CC PIR; F83188; F83188.
CC PIR; JC4832; JC4832.
CC InterPro; IPR000374; PC trans.
CC Pfam; PF01148; CTP trans1.1.
CC PROSITE; PS01315; CDS; 1.
CC Transferrase; Nucleoside transferase; Phospholipid biosynthesis;
CC Transmembrane; Inner membrane. Complete proteome.
CC TRANSMEM 12 32
CC TRANSMEM 53 73 POTENTIAL.
CC TRANSMEM 75 95 POTENTIAL.
CC TRANSMEM 111 131 POTENTIAL.
CC TRANSMEM 136 156 POTENTIAL.
CC TRANSMEM 174 194 POTENTIAL.
CC TRANSMEM 199 219 POTENTIAL.
CC TRANSMEM 251 271 POTENTIAL.
CC TRANSMEM 131 133 WPL -> VAA (IN REF. 1).
CC CONFLICT 131 133
CC SEQUENCE 271 AA; 28856 MW; 5025059C3F1A64C7 C5C64;
Query Match 19.1%; Score 298; DB 1; Length 271;
Best Local Similarity 33.3%; Pred. No. 1.2e-15;

```

```

Matches 98; Conservative 38; Mismatches 94; Indels 64; Gaps 15;
QY 22 RDLKRAIAVIGIGLVLL-----CIVLS--PWGMYILVAGF-----MAAATWE 63
DB 5 RIITALLVLPALGGFFLEGFAFFALFICAVVSLGAWEARL-AGVEQFGVAVAAVA 63
QY 64 VSRLEKGGYHLPPLMIIGQAIITWSPFGTGMILASFVATVLMVYFRFYNGTEKE 123
DB 64 V---LMVALYHLP---QAGAVLLALVW-----WTLATVILVLT-----PE 99
QY 124 ARNYL--RDTSGVIFVLTWIPFGSFAAMLMLQNSNIPGTYFILTFLMCLVIAVDGGYI 181
DB 100 SVGYWGRWRRLGMGLILLPAQGLVLLKQWPLANGL-----IAVVLVWGADIGAVF 154
QY 182 AGVFGSHPMAPLVSPKSWEGPAGSIVLG-SVTGALSVHFLLDHHWVG-----VILGA 236
DB 155 SGKAFGRKRLAPRVSPKSGWEGVYGLAASLAITLAVGLY-----RGWSLGALLALLGAA 210
QY 237 LVVCAVL-GDLVESQFKDLGIKDKSNLPGHGGMLDRLDGLMPLA-----AMVTW 285
DB 211 LVFVSIVGDLTESMFKRSGINDSNLLPGHGGVLDRLDSTLTAIPVPAALLW 264
RESULT 4
CDS: ECOLI
ID CDS: ECOLI STANDARD; PRT: 249 AA.
AC P06466;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
DE synthase) (CDS) (CTP-phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
GN CDS OR CDS OR B0175 OR Z0186 OR ECS0177 OR SF0165 OR S0168.
OS Escherichia coli,
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=86008268; PubMed=2995358;
RA Icho T., Sparrow C.P., Raetz C.R.H.;
RT "Molecular cloning and sequencing of the gene for CDP-diglyceride
RT synthetase of Escherichia coli.";
RL J. Biol. Chem. 260:12078-12083(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=KL2 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=KL2 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,

```

RA Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.,
RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=0157:H7 / EBL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perrin N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Bourin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potancusis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=0157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:111-22 (2001).
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
RN [9]
RP CHARACTERIZATION.
RC SPECIES=E.coli;
RX MEDLINE=86008269; PubMed=2995359;
RA Sparrow C.P., Raetz C.R.H.;
RT "Purification and properties of the membrane-bound CDP-diglyceride
RT synthetase from Escherichia coli.";
RL J. Biol. Chem. 260:12084-12091 (1985).
CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
CC diacylglycerol.
CC -!- PATHWAY: Phospholipid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: Belongs to the CDS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11330; AAA23545.1; -;
DR EMBL; AE000127; AAC73286.1; -;

DR EMBL; D83536; BAA77850.1; -;
DR EMBL; U70214; AAB08604.1; -;
DR EMBL; AE005193; AAG54477.1; -;
DR EMBL; AP002550; BAB33600.1; -;
DR EMBL; AE015053; AAN41827.1; ALT_INIT.
DR EMBL; AE016978; AAP15708.1; -;
DR PIR; A23898; SYECDG.
DR PIR; A85502; A85502.
DR PIR; A99651; A99651.
DR EcoGene; EGI0139; cdaA.
DR InterPro; IPR000374; PC trans.
DR Pfam; PF01148; CTP_transf_1; 1.
DR PROSITE; PS01315; CDS; 1.
KW Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;
KW Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 57 77 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 115 135 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 228 248 POTENTIAL.
SQ SEQUENCE 249 AA; 27571 MW; F628EC847F193647 CRC64;
Query Match 18.2%; Score 285; DB 1; Length 249;
Best Local Similarity 31.0%; Pred. No. 18-14; Indels 64; Gaps 12;
Matches 85; Conservative 38; Mismatches 87;
QY 47 WG-----WYILVAGEMAATMEVGRLEGGYHLPIMIGQAIWLS--- 91
DB 7 WQQLSGFTTRSORVMLAVLGLLMLFL---LPEYHNIHQPLVEIS---LWASLW 59
QY 92 WPGTGMILASVAIVLVMYFRIFNGTEKARNYLRDTSVGIF-VLTWPLF-GSEAA 149
DB 60 W-----IVALLVL-----FYPGSAATWRN--SKTLIFGVITVFFWGMAL 102
QY 150 MSLMQNNSIPGTFTPLMCLVIASDVGGYIAGVFFGSHPNAPLVSPKSWEGFAGSIV 209
DB 103 RAWHYDENHYSGAIVLLYVILWVGADSGAYWFGKLFGRKLPKVPKTKWQGFIGGLA 162
QY 210 LGSVTGALSVHFLDHHWNGV-----ILGALV--VCATLGLDVSQFKRDGLI 257
DB 163 TRAVIS-----WGYNWALNDVAPVTLTLCISVAALASVIGDLTSMFKREAGI 211
QY 258 KQMSNLLPGHGLMDRLDGLPAAMVTLILSVI 291
DB 212 KQSHLIPGHGILDRIDSLTRAAPVFACLLLV 245
RESULT 5
CDS_A HAEIN STANDARD; PRT; 288 AA.
AC P44937; Q32623; Q32627;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
DE CDS OR CDS OR H10919.
GN Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
[1]
SEQUENCE FROM N.A.
STRAIN=rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness B.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

Db 216 TLVGRNITGFIILSVATVAISVLGDLITESTMFKRESGVKDSQLIPHGCVLDRIDSLTA 275
 Qy 280 A 280
 Db 276 A 276
 RESULT 6
 CDSA AQUAE STANDARD; PRT; 259 AA.
 ID CDSA AQUAE
 AC 067292;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphatidate cyclyglyltransferase (EC 2.7.7.41) (CDP-diglyceride
 synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
 synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
 synthase) (CDP-DG synthetase).
 DE CDSA OR CDS OR AQ_1249.
 OS Aquifex aeolicus
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxId=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus";
 RT Nature 392:353-358(1998).
 RL Nature 392:353-358(1998).
 CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
 diacylglycerol.
 CC
 CC -!- PATHWAY: Phospholipid biosynthesis.
 CC
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the CDS family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announcement>)
 or send an email to license@sb-sib.ch.
 CC
 CC EMBL; AE000730; AAC07246.1; -.
 CC PIR; A70408; A70408.
 DR InterPro; IPR000374; PC_trans.
 DR Pfam; PF01148; CTP_transf_1; 1.
 DR PROSITE; PS01315; CDS; 1.
 DR Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;
 KW Transmembrane; Complete proteome.
 FT Transmembrane 31 51 POTENTIAL.
 FT TRANSMEM 69 89 POTENTIAL.
 FT TRANSMEM 103 123 POTENTIAL.
 FT TRANSMEM 129 149 POTENTIAL.
 FT TRANSMEM 170 190 POTENTIAL.
 FT TRANSMEM 193 213 POTENTIAL.
 FT TRANSMEM 236 256 POTENTIAL.
 SQ SEQUENCE 259 AA; 29486 MW; D5FB386647BD0D0E CRC64;
 Query Match 16.1%; Score 251.5; DB 1; Length 259;
 Best Local Similarity 29.0%; Pred. No. 3.6e-12;
 Matches 80; Conservative 58; Mismatches 101; Indels 37; Gaps 11
 Qy 17 KKNAGRD:KAAIAGVIGLGVLLGVLSPWGWYLVAGFMA-AAWVEYGSLSKEGGYHL 75
 Db 8 RSSGEFLMSREFGVGLIGTTL-VIFLPKSLFLVLVILFLCFA-SREVSVALGENEVFY 66
 Qy 76 PLPIMIGQAIWLWNP--FGTWGILASVATVLMVYFRIFYNGTEKEARNYLRDTSV 133
 Db 67 RSPVLV----TYFADPLVFPVPLGLISLVA-----YKR-----WELNSFFKST-- 107

140	WT	-----PFGSF-----	RAMLSLMON	156
:	:	:	:	:
63	WLI	VTGILLILDRGALLTIGFLVAGCAILLV	TQWKSGRGWPAAGLFYAGFSALSLLRG	122
:	:	:	:	:
157	NSI	PGTFIITFMLCVI-ASDGGYIAGVFFG	SHPMAPLSPVKSWEGFAGSIVLGSVTG	215
:	:	:	:	:
123	DE	-PFGTTIVFLFAVVMSTDI	TAYFNRRALGQPKJAPFESNKTWSGAIGAAAAVAGG	181
:	:	:	:	:
216	ALS	VFELDHHHMMGVTILGCALVVCATLGL	DLVBSQFKRDLGIDKMSNLLPGHGGLMDRLD	275
:	:	:	:	:
192	LIV	ASLVAAACGGGVPIALLLSIVSQIGD	LAESWYKQFGAKDSGLLPGHGGVLDRVD	241
:	:	:	:	:
276	GML	PAAAVTWL-----ILSVISSY	295	
:	:	:	:	:
242	GLV	AAALLYFGAIFAEPPDVLGAIFFSF	270	
:	:	:	:	:

DB 242 GLVAAALLYLFGAIPAEPAEPOVLGAIFFSF 270

RESULT 8

	CDSA BRUSU	STANDARD;	PRT;	270 AA.
ID	CDSA BRUSU			
AC	Q8GOE0;			
DT	15-MAR-2004	(Rel. 43, Created)		
DT	15-MAR-2004	(Rel. 43, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG synthase) (CDP-DG synthetase).			
DE	CDSA OR BR1157.			
GN	Bruceella suis.			
OS	Bruceella suis.			
OC	Bacteriellae; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Bruceellaceae; Bruceella.			
OX	NCBI_TaxID=29461;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RY	STRAIN=1330 / Biovar 1;			
RX	MEDLINE=22247741; PubMed=122711122;			
RA	Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F., Read T.D., Dodson R.J., Unayam L.R., Brinkac L.M., Beanan M.J., Daugherty S.C., Deboy B.T., Durkin A.S., Kolonay J.F., Madupu R., Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E., Riedmuller S., Tattalin H., Gill S.R., White O., Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.; "The Bruceella suis genome reveals fundamental similarities between animal and plant pathogens and symbionts.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).			
RL	-!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-diacylglycerol.			
CC	-!- PATHWAY: Phospholipid biosynthesis.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (By similarity).			
CC	-!- SIMILARITY: Belongs to the CDS family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AE014416; AAN30077.1; -.			
DR	TIGR; BR1157; -.			
DR	InterPro; IPR000374; PC trans.			
DR	Pfam; PF01148; CTP transf_1; 1.			
DR	PROSITE; PS01315; CDS; 1.			
KW	Transferase; Nucleotidyltransferase; Phospholipid biosynthesis;			
FT	Transmembrane; inner membrane.			
FT	TRANSMEM 19 39 POTENTIAL.			
FT	TRANSMEM 53 73 POTENTIAL.			
FT	TRANSMEM 76 96 POTENTIAL.			
FT	TRANSMEM 101 121 POTENTIAL.			
FT	TRANSMEM 126 146 POTENTIAL.			

DR	PROSITE; PS03315; CDS; 1_
KW	Transferase; Nucleotidyltransferase; phospholipid biosynthesis;
KW	Transmembrane; Inner membrane.
FT	Transmemmem 39 POTENTIAL.
FT	Transmemmem 53 POTENTIAL.
FT	Transmemmem 76 POTENTIAL.
FT	Transmemmem 96 POTENTIAL.
FT	Transmemmem 101 POTENTIAL.
FT	Transmemmem 121 POTENTIAL.
FT	Transmemmem 126 POTENTIAL.
FT	Transmemmem 146 POTENTIAL.

DE synthase) (CDP-DG synthetase).
GN CDSA OR HP0215.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Keriavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori";
RL Nature 388:539-547(1997).
CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
diacylglycerol.
CC -!- PATHWAY: Phospholipid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(BY similarity).
CC -!- SIMILARITY: Belongs to the CDS family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000541; AAD07283.1; -.
DR PIR; G64546; G64546.
DR TIGR; HP0215; -.
DR InterPro; IPR000374; PC trans.
DR Pfam; PF01148; CTP transf. 1;
DR PROSITE; PS01315; CDS; 1;
DR Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;
KW Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 78 98 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.
FT TRANSMEM 125 145 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
SQ SEQUENCE 266 AA; 28899 MW; 4995F8D1B92904F4 CRC64;
Query Match 15.3%; Score 239.5; DB 1; Length 266;
Best Local Similarity 30.2%; Pred. No. 2.9e-11;
Matches 68; Conservative 28; Mismatches 90; Indels 39; Gaps 5;
QY 81 IIGQAIILWSPFGTGMILASP-VATVLVLMYFRIFYNGT----- 120
DB 37 VLGGIYAVGFSEALRLFOVKASFSLYLILVLSWAAYFNGRPIECALISAMVAVIAYQ 96
QY 121 ---EKEARNYLDTSVGLFVLTPFLGFSFAMLSLQNNISIPGYTILTMCLVIAADV 177
DB 97 KAHSEAILPFLYPGVGFFA-----LFGVYKDF-----GAVAILLWLVVVYASDV 141
QY 178 GGYTAGVFGGHPMAPLVSPKKSMEGFAGSVLGSVTVGALSVMHFLDLHHMMGVILCAL 237
DB 142 GAFFGKLLGKTPTFP-TSPNKTLEGALIGVVLASVLGSFVCMGKLSGGFFNALFFSFLI 200
QY 238 VVCATLGLVFSQKRDIGIKDMSNLLPFGHGLMDRLDGLMPLAAM 282

DB 201 ALVAVFGDLYESYLKRVKVGIKDSKILPGHGVLDRLDSMLFGAL 245
RESULT 11
CDSA_HELPJ STANDARD; PRT; 266 AA.
ID Q9ZML7;
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
GN CDSA OR JHP0201.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nikelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori";
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
diacylglycerol.
CC -!- PATHWAY: Phospholipid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(BY similarity).
CC -!- SIMILARITY: Belongs to the CDS family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001459; AAD05785.1; -.
DR PIR; F71961; F71961.
DR InterPro; IPR000374; PC trans.
DR Pfam; PF01148; CTP transf. 1;
DR PROSITE; PS01315; CDS; 1;
DR Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;
KW Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 78 98 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.
FT TRANSMEM 125 145 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
SQ SEQUENCE 266 AA; 28788 MW; 99C48ED30B45AF45 CRC64;
Query Match 14.8%; Score 231.5; DB 1; Length 266;
Best Local Similarity 28.8%; Pred. No. 1.2e-10;
Matches 77; Conservative 35; Mismatches 110; Indels 45; Gaps 8;
QY 39 LIGIVLSPGWITLVAGFMAAT-----NEVGSRLKEGGYHLPIMIGG-----AIIW 89
DB 13 ITGWL-----IVVAGLILYADNLLFWALVGGIYAVGFSEALRLFOVKASFSLYLIV 66
QY 90 LSN-----PFGTMGILASFVATVLVLMYFRIFYNGTKEARNYLDTSVGFVLTWI 141
DB 67 LSWVAAFYNGHVEFCALISAMVAVIA-----YKAHSEAILPFLYPGVGFFA----- 116

QY 142 PLFGSFAAMLSMNNQNSIPGTYFILTFLMVCVIAADGVGTYAGVFFGSHPMAPLVSPKSW 201
 Db 117 -LFGVYKDF-----GAVALIWLVLVVVADVGAFGGKGLGTPTT-ATSPNKL 164
 QY 202 EGFAGSIVLGSVTGALSVEHLLDHHWGMVILGALVVCATLGLVDSQFKRDIGIKDMS 261
 Db 165 EGALLGVVLASVLSFGVMGKLSGFLMALFLSFLIALMAVFGDLYESYLKRVGVKDSG 224
 QY 262 NLLPCHGGMLDRDLCMLPAAMVTWLIL 288
 Db 225 KILPCHGGVLDRLDLSMLFGALSLHVL 251

RESULT 12

YNBB_ECOLI STANDARD; PRT; 298 AA.

AC P76091; P78229;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein ynbB.
 GN YNB5 OR B1409.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474 (1997).
 RN [2]

SEQUENCE OF 187-298 FROM N.A.

RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishio Y., Oshima T., Saito N.,
 RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377 (1996).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the CDS family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB000238; AAC74491.1; -;
 DR EMBL; D90779; BAAL5023.1; -;
 DR EMBL; D90780; BAAL5027.1; -;
 DR PIR; D64892; D64892.
 DR EcoGene; EG13749; ynbB.
 DR InterPro; IPR000374; PC trans.
 DR Pfam; PF01148; CTP transf.1; 1.
 DR PROSITE; PS01315; CDS; 1.
 DR Hypothetical protein; Transmembrane; Transferase; Complete proteome.

FT TRANSMEM 5 25
 FT TRANSMEM 52 72
 FT TRANSMEM 105 125

FT TRANSMEM 138 158 POTENTIAL.
 FT TRANSMEM 163 183 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 236 256 POTENTIAL.
 FT TRANSMEM 273 293 POTENTIAL.
 FT TRANSMEM 287 307 W -> M (IN REF. 2).
 FT CONFLICT 187 197 W -> M (IN REF. 2).
 FT CONFLICT 190 199 S -> I (IN REF. 2).
 SQ SEQUENCE 298 AA; 33053 MW; 8083BA7E3C07FADF CRC64;

Query Match 14.7%; Score 230.5; DB 1; Length 298;

Best Local Similarity 26.1%; Pred. NO. 1.6e-10;
 Matches 76; Conservative 47; Mismatches 121; Indels 47; Gaps 11;

QY 12 RMPKPKNNAGRDIAIAVIGIGLVILGIVLSPGW---YIIVAGFMAAATWEVGSR 67
 Db 26 RLPERRGG-----EVLRIRTWGVICFMSWISGRWMTLTFALISFLA----- 71
 QY 68 LKEG-----GYHLPLPIMIIGQAIIWLSPFGTNGILASFVATVLVLMY-----PRIFYN 118
 Db 72 LKEYCTLSVHPFR-----WLYGIPLNVLLIGFNCFFELFLFIPLAGFLILAT 120
 QY 119 G--TEKEARNYLRDTSVGIF---VLTWIFLFGSFAAMLSMNNQNSIPGTYFILTFLMVCV 173
 Db 121 GQVLVGDPSGFLHTVS-AIFWGWIMTFAL--SHAALLMLPTTNIQGGALLVFLILALT 177
 QY 174 AS-DVGGYIAGVFFGSHPMAPLVSPKSWEGFAGSIVLGSVTGALSVEHLLDHHWGMV 232
 Db 178 ESNDIAQYLWGKSCGRKVKVPKVTLEGMLGGVITIMIASLIIGLPLTLNTLQALL 237
 QY 233 LGCALVVCATLGLVDSQFKRDIGKDMNSLLPCHGGMLDRDLCMLPAAMV 283
 Db 238 AGLLIGISGFCGVDVVMISAIRDIGNKDSGKLLPGHGLDRIDSIFTAPV 288

RESULT 13

CDSA_SINY3 STANDARD; PRT; 293 AA.

AC P73548;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
 DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
 DE synthase) (CDS) (CTP-phosphatidate cytidyltransferase) (CDP-DAG
 GN CDSA OR SLR1369.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose M., Sugita M., Saito M., Saito M., Saito M.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimizu S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136 (1996).
 CC -!- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
 CC diacylglycerol.
 CC -!- PATHWAY: Phospholipid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the CDS family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```

or send an email to license@isb-sib.ch).
-----
CC EMBL; D90907; BAA17588.1; -.
CC PIR; S77254; S77254.
CC InterPro; IPR000374; PC trans.
CC Pfam; PF01148; CTP trans1_1; 1.
CC PROSITE; PS01315; CDS; 1.
CC Transmembrane; Nucleotidyltransferase; Phospholipid biosynthesis;
KW Transmembrane; Complete proteome.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 97 117 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 195 215 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 273 293 POTENTIAL.
SQ SEQUENCE 293 AA; 31875 MW; A/E3ABC9C17475D0 CRC64;

Query Match 14.3%; Score 224; DB 1; Length 293;
Best Local Similarity 27.9%; Pred. No. 4.7e-10;
Matches 87; Conservative 49; Mismatches 112; Indels 64; Gaps 16;

22 RDLKAAIVGIGLVVLGVLGVLSPGWYIIVAGPMAAATWVGSRLKE-----70
Db 5 RIISAVIGIALAFSLIL-----GGWY-----FSAAILVIYLGUREYFQWVRAGIA 52
QY 71 --GGVHLPLPIMIGQAIWLSWP-----FGTGILASF-----VATV--LVLMYF 113
Db 53 PAAKTTWLSLMLLSAIVT-----PHLDAFPPLTGALICVYLLFQPKMATIADISTLL 108
QY 114 RIFYNG--TEKEARNYLRDTSVIGFVLVPLFGSFAAMLSLQNNISIPGYFYLTWMLC 171
Db 109 GLFYGYGLPSYVWRLRGDGVNPMGL--HLPLNGFWPESWAHPNFPTGLLVTLAF--AC 166
QY 172 VTASDVGVYAGVFGFHPMAPLVSPKSWE-----GFAGSIVLGSVGTGALSVEHFLDHW 227
Db 167 IWAADIGAYTMKGWGRTRLS--ISPKATVSGSLVGVGSLLVG--VLGA-----WYLOPY 220
QY 228 W--MGVILGALVCAVATGLDVLVESQFRLDGKIDMSNLLPGHGLMDRLDGLMPLAAMVTW 285
Db 221 WEITGALLGLLIGVSLGLDLSMTMKRDAGVDSGQLIPGHGGILDRTSDSYVFTAPLVY 280
QY 286 ----LILSVISS 293
Db 281 YFVILLPLVNN 292

RESULT 14
CDS: RICTP STANDARD; PRT; 228 AA.
AC Q9ZDAB;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
GN CDS OR RP424.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=782;
OX [1]
RN R1
RP SEQUENCE FROM N.A.
RX STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichertitz-Ponten T., Almqvist U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."

```

```

RL Nature 396:133-140(1998).
CC -I- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
CC diacylglycerol.
CC -I- PATHWAY: Phospholipid biosynthesis.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -I- SIMILARITY: Belongs to the CDS family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ235271; CAA14881.1; -.
CC PIR; G71700; G71700.
CC InterPro; IPR000374; PC trans.
CC Pfam; PF01148; CTP trans1_1; 1.
CC PROSITE; PS01315; CDS; 1.
CC Transmembrane; Nucleotidyltransferase; Phospholipid biosynthesis;
KW Transmembrane; Complete proteome.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
FT TRANSMEM 228 AA; 25814 MW; 48DE66AB6EA19174 CRC64;
SQ SEQUENCE 228 AA; 25814 MW; 48DE66AB6EA19174 CRC64;

Query Match 13.9%; Score 216.5; DB 1; Length 228;
Best Local Similarity 29.1%; Pred. No. 1.4e-09;
Matches 67; Conservative 41; Mismatches 77; Indels 45; Gaps 11;

QY 79 IMIIGQAIWLSWPFGTWGLASFVATV-----VLMYFRI-----FYNGTEKAR 125
Db 19 LRIISGIALVSL-----FVIALCLKTLFYILMILVGLGMLSEWNYMTPSI- 65
QY 126 NYLRDTSVIGFVLVPLFGSFAAMLSLQNNISIPGYFYLTWMLCVIAS-DVGVIAGV 184
Db 66 NYL---LIGLII---IPIPSILLIFLSMEESNRL-----VIMLYFCILWSVDTFAMIGK 114
QY 185 FPGSHPMALVSPKSWEG-FAGSIVLG--SVTGALSVEHFLDHWG-----VILGAL 237
Db 115 TFKGIKLAPKISPKTWTGLITGTVSAGLVSVLSIPNYHIEHYFNSKIVLFIISIL 174
QY 238 VVCAITGLDVLVESQFRLDGKIDMSNLLPGHGLMDRLDGLMPLAAMVTWLI 287
Db 175 ALIAQSSDLFISVFKRKFNKIDSGHIIPGHGGVLDKDFDSIILTAPVFFCI 224

RESULT 15
CDS: RICTP STANDARD; PRT; 270 AA.
AC Q9X1B7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
DE synthase) (CDP-DG synthetase).
GN CDS OR TM1397.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
OX [1]
RN R1
RP SEQUENCE FROM N.A.
RX STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickley E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,

```

Search completed: August 11, 2004, 13:35:09
Job time : 15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:32:07 ; Search time 39 Seconds
(without alignments)
2402.792 Million cell updates/sec

Title: US-09-853-641-2
Perfect score: 1563
Sequence: 1 MNEPEQHRSNRMKPKKNA.....LPAAVWTLILSVISSYPS 297

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phase:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_rvirus:**
- 16: sp_bacteriap:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1495	95.6	285	16 Q8NP04	Q8NP04 corynebacte
2	1366	87.4	297	16 Q8FP76	Q8FP76 corynebacte
3	551	35.3	392	16 Q82JY1	Q82JY1 streptomyc
4	539	34.5	391	16 Q86769	Q86769 streptomyc
5	350.5	22.4	328	16 Q8G482	Q8G482 bifidobacte
6	322.5	20.6	280	16 Q83G74	Q83G74 tropheryma
7	322	20.6	269	16 Q83NN2	Q83NN2 tropheryma
8	310	19.8	271	16 Q885N8	Q885N8 pseudomonas
9	310	19.8	288	16 Q7VRE0	Q7VRE0 candidatus
10	299	19.1	271	16 Q88MH5	Q88MH5 pseudomonas
11	299	19.1	282	16 Q8ZH60	Q8ZH60 yersinia pe
12	288.5	19.1	289	16 Q7VM21	Q7VM21 haemophilus
13	298	19.1	280	16 Q9KPV7	Q9KPV7 vibrio chol
14	298	19.1	285	16 Q8XG59	Q8XG59 salmonella
15	296	18.9	285	16 Q8FL09	Q8FL09 escherichia
16	291	18.6	272	16 Q83BV6	Q83BV6 coxiella bu

17	286.5	18.3	265	16 Q9JX34	Q9JX34 neisseria m
18	282	18.0	275	16 Q8PML0	Q8PML0 xanthomonas
19	280.5	17.9	265	16 Q9K1G7	Q9K1G7 neisseria m
20	280.5	17.9	289	16 Q9CJL3	Q9CJL3 pasteurella
21	279	17.9	275	16 Q8PAV8	Q8PAV8 xanthomonas
22	278	17.8	280	16 Q8A0L5	Q8A0L5 bacteroides
23	278	17.8	280	16 Q8A0L5	Q8A0L5 bradyrhizob
24	276.5	17.7	285	16 Q7TU77	Q7TU77 prochloroco
25	274.5	17.6	273	16 Q8KFT9	Q8KFT9 chlorobium
26	272	17.4	282	16 Q8DBF6	Q8DBF6 vibrio vuln
27	271	17.3	280	16 Q87ME2	Q87ME2 vibrio para
28	268	17.1	271	16 Q8XZ16	Q8XZ16 ralstonia s
29	265.5	17.0	294	16 Q8YQF6	Q8YQF6 anabaena sp
30	265	17.0	348	16 Q7USC6	Q7USC6 rhodospirell
31	259.5	16.6	284	16 Q9RU83	Q9RU83 deinococcus
32	257.5	16.5	296	16 Q7VBI1	Q7VBI1 prochloroco
33	256.5	16.4	277	16 Q92Q50	Q92Q50 rhizobium m
34	254	16.3	298	16 Q7U7N9	Q7U7N9 synechococc
35	253	16.2	265	16 Q895K7	Q895K7 clostridium
36	253	16.2	399	10 Q9M001	Q9M001 arabidopsis
37	252.5	16.2	391	10 Q9A003	Q9A003 arabidopsis
38	252.5	16.2	430	10 Q7XJM7	Q7XJM7 arabidopsis
39	252	16.1	309	16 Q87TX6	Q87TX6 pseudomonas
40	251	16.1	310	16 Q8PQ11	Q8PQ11 xanthomonas
41	250.5	16.0	264	16 Q8XJRO	Q8XJRO clostridium
42	250	16.0	304	16 Q7TUW2	Q7TUW2 prochloroco
43	249	15.9	241	16 Q8PMW2	Q8PMW2 campylobact
44	248.5	15.9	278	16 Q8YHH2	Q8YHH2 brucella me
45	247.5	15.8	260	16 Q8CST9	Q8CST9 staphylococ

ALIGNMENTS

RESULT 1

Q8NP04	Q8NP04	PRELIMINARY;	PRT;	285 AA.
AC	Q8NP04;			
DT	01-OCT-2002 (TREMELrel. 22, Created)			
DT	01-OCT-2002 (TREMELrel. 22, Last sequence update)			
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)			
DE	CDP-diglyceride synthetase (EC 2.7.7.41).			
GN	CGL2022.			
OS	Corynebacterium glutamicum (Brevibacterium flavum).			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.			
OX	NCBI_TaxID:1718;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;			
RA	Nakagawa S.;			
RT	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";			
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP005280; BAB99415.1; -			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.			
DR	GO; GO:0016740; P:transferase activity; IEA.			
DR	GO; GO:0008654; P:phospholipid biosynthesis; IEA.			
DR	InterPro; IPR000374; PC trans.			
DR	Pfam; PF01148; CTP transf_1; 1.			
DR	PROSITE; PS01315; CDS; 1.			
KW	Transferase; Nucleotidyltransferase; Complete proteome.			
SQ	SEQUENCE 285 AA; 30733 MW; E41025439D7002B4 CRC64;			

Query Match 95.6%; Score 1495; DB 16; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.5e-107;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 MPKPKNNAGRDLLKAAIAVGIGLVLLGIVLSPWGVILVAGFMAATWVGSRLEKGG 72

Db 1 MPKPKNNAGRDLLKAAIAVGIGLVLLGIVLSPWGVILVAGFMAATWVGSRLEKGG 60

OY 73 YHLPUPIMIGQAIWLSPFGTGMGILASFVATVILVLMYRIFVNGTEKEARNYLRTS 132

```

Db 61 YHPLPIMIGQALIWSPFGTGMGILASVATVVLVLMVFRIFYNGTEKEARNYLDTS 120
QY 133 VGIFVLTWIPFGSPAAMLSLMONNSIPGTFFILTMCLVITASDVGGYIAGVFGSHPM 192
Db 121 VGIFVLTWIPFGSPAAMLSLMONNSIPGTFFILTMCLVITASDVGGYIAGVFGSHPM 180
QY 193 PLVSPKSWEGFAGSIVLGSVTGALSVDHLDHWMGMVILGCAVVCATIGDLVSGQFK 252
Db 181 PLVSPKSWEGFAGSIVLGSVTGALSVDHLDHWMGMVILGCAVVCATIGDLVSGQFK 240
QY 253 RDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVISSYPS 297
Db 241 RDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVISSYPS 285

RESULT 2
Q8FP76 PRELIMINARY; PRT; 297 AA.
AC Q8FP76;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative phosphatidate cytidyltransferase.
GN CDS OR SAV2623.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR ENBL; AP005031; BAC70334.1;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000374; PC trans.
DR Pfam; PF01148; CTP transf_1; 1.
DR PROSITE; PS01315; CDS; 1.
DR Nucleotide transferase; Transferase; Complete proteome.
KW Nucleotide transferase; Transferase; Complete proteome.
SQ SEQUENCE 297 AA; 32053 MW; 634F0906B4E06915 CRC64;

Query Match 87.4%; Score 1366; DB 16; Length 297;
Best Local Similarity 83.7%; Pred. No. 1.3e-97;
Matches 247; Conservative 30; Mismatches 18; Indels 0; Gaps 0;

QY 1 MNEPEQHRSRMMPKPKNAGRDLLKAAIATVGLGVLVLLGIVLSPWGWYILVAGFMAA 60
Db 1 MSEPHEHRTVMPKPKNSAGSRLKAAVGVGVGLLVIVGILLSPGWYILVAGAVAVA 60

QY 61 TWEVGSRLKEGYHPLPLTMIIGQAIWLSPFGTGMGILASVATVVLVLMVFRIFYNGT 120
Db 61 TWEVCGRLKEAGYHPLPMPIMIIGQAVWCSPFFETSGILASYATVVLVLMFFRIFYNGP 120

QY 121 EKEARNYLDTSVGFVLTWIPFGSPAAMLSLMONNSIPGTFFILTMCLVITASDVGGY 180
Db 121 EKEARNYLDTSVGFVLTWIPFGSPAAMLSLMONNSIPGTFFILTMCLVITASDVGGY 180

QY 181 IAGVFGSHPMAPLVSPKSWEGFAGSIVLGSVTGALSVDHLDHWMGMVILGCAVVC 240
Db 181 AAGVFFGAHPMAPVSPKSWEGFAGSIVLGSLTGATITVHLLHHHWWLGVILGILVVC 240

QY 241 ATIGDLVSGQFKRDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVISSY 295
Db 241 ATIGDLVSGQFKRDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVISSY 295

```

```

RESULT 3
Q82JY1 PRELIMINARY; PRT; 392 AA.
AC Q82JY1;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative phosphatidate cytidyltransferase.
GN CDS OR SAV2623.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR ENBL; AP005031; BAC70334.1;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000374; PC trans.
DR Pfam; PF01148; CTP transf_1; 1.
DR PROSITE; PS01315; CDS; 1.
DR Nucleotide transferase; Transferase; Complete proteome.
KW Nucleotide transferase; Transferase; Complete proteome.
SQ SEQUENCE 392 AA; 40876 MW; C6F6A4254E7C8AC6 CRC64;

Query Match 35.3%; Score 551; DB 16; Length 392;
Best Local Similarity 39.1%; Pred. No. 1.3e-34;
Matches 117; Conservative 52; Mismatches 114; Indels 16; Gaps 5;

QY 3 EPEQHRSRMMPKP-----KNNAGRDLLKAAIATVGLGVLVLLGIVLSPWGWYILVAGFM 57
Db 102 QPKPEPMPDAPQAPAPQPKSAGRDLLGAAIGVGVGLVIVASLVFVKAVFVGIVAV 161

QY 58 AATWVGSRLEKE-GGYHPLPLTMIIGQAIWLSPFGTGMGILASVATVVLVLMVFRIF 116
Db 162 VGLWELTSLERLBERKIKAPLVPLAVGGAAVYVAGVYRGAEGAVAMALTALAVLWRM- 220

QY 117 YNGTEKEARNYLDTSVGFVLTWIPFGSPAAMLSLMONNSIPGTFFILTMCLVITASD 176
Db 221 ----TEPPEGYLRDVTAGVFAAFYVFPFLATFVAMMLTADD----GPRRYLTLTLTVSD 272

QY 177 VGGYIAGVFGSHPMAPLVSPKSWEGFAGSIVLGSVTGALSVDHLDH-HWMMGVILQC 235
Db 273 TGAATAGWRFGRHRLAPRISPKRTREGLGVAGSVFANVAGALCMFEFLIDDSWQGVILGF 332

QY 236 ALVVCATLGDVSGQFKRDLGKIDMSNLLPGHGLMDRLDGMPLPAAMVTWILSVISS 294
Db 333 AVAASATLGDGSSMIKRDLGKIDMGTLIPGRGGIMDRDLSLLPTAPVYVWLLIVFVS 391

RESULT 4
Q86769 PRELIMINARY; PRT; 391 AA.
ID Q86769
AC Q86769;

```


Db 278 VQWRFGRKKLAPRISPCKTREGLLGATAFANVAGALCMQFLIDDGAWQGLLLGLTVAVS 333

QY 241 ATFLGLVSESQFKDLGIGKMSNLLPGRGGLMDRLDGMPLPAAMVTWILSVISS 294

Db 338 ATFLGLGESSMIKDLGIGKDMGTLLPGRGGIMDRLDLSLLPTAPVVWLLLVIFVGS 391

RESULT 5

Q8G482 PRELIMINARY; PRT; 328 AA.

ID Q8G482; AC Q8G482; DT 01-WAR-2003 (TREMBlrel. 23, Created) DT 01-WAR-2003 (TREMBlrel. 23, Last sequence update) DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update) DE Probable phosphatidate cytidyltransferase. CDSA OR BL1507. OS Bifidobacterium longum. OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales; CC Bifidobacteriaceae; Bifidobacterium. NCBI_TaxID=216816; [1]_RN SEQUENCE FROM N.A. RP SEQUENCE FROM N.A. RC STRAIN=NCC 2705; RX MEDLINE=22294377; PubMed=12381787; RA Schell M.A., Karimiantzou M., Snel B., Vilanova D., Berger B., Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M., Pridmore R.D., Arigoni F.; RT "The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract."; RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002). RL EMBL; AF014782; AAN25302.1; -. DR GO; GO:0016020; C:membrane; IEA. DR GO; GO:0004605; F:phosphatidate cytidyltransferase activity; IEA. DR GO; GO:0016740; F:transferase activity; IEA. DR GO; GO:0008654; P:phospholipid biosynthesis; IEA. DR InterPro; IPR000374; PC:trans. DR Pfam; Pf01148; CIP_transf1; I. DR Nucleotide transferase; Transferase; Complete proteome. KW SEQUENCE 328 AA; 35279 MW; B45CBA105FC37839 CRC64;

Query Match 22.4%; Score 350.5; DB 16; Length 328; Best Local Similarity 31.0%; Pred. No. 3.1e-19; Matches 104; Conservative 52; Mismatches 119; Indels 61; Gaps 12

QY 1 MNEPQHERSMR--MPKPKNNAGRDLLKAAIAGVIGLVLVILGIVLSPGWVILVAGFWA 58

Db 1 MERNEQPEHEAEETLDQINKTKTGRNMPQAIATGAALVILIVACLLSVDLVLVVVFEI 60

QY 59 RATVEVGSRLKEGGVHLPLPIMIIGQAIIWLSWPG-----TMG-----I 99

Db 61 LALWEIRVDFPATVGHLPV-----FMWLCCSFTLIATYSPHLITNGLSIIVSI 112

QY 100 LASFVATVVLVIMYFRIF-----YNGTEKEAR-----NYLRDTSVGIFVLT-- 139

Db 113 LVIAAASAKISVGNRLSLAVAGKLSNTDASARLESSFNHGGSQHHSLSHVAVSVLTVL 172

QY 140 WTLPGSFAAMLMLMQNNISPGYIFILTFMLCVIADVGVIAGVFFGSHPMAPLVSPKK 199

Db 173 YIPLDAS--CVIISFTFNHGPFAHAIMLVFPLP--SDTGGLFAGAWLGHKHSPLRISPKK 229

QY 200 SWEFGPAGSIVLGSVTGALSVEHFLDHF-----HWNMGVILGALVVCATLDGLVSESQFK 252

Db 230 SVEGLVGS-MLFAMAGAFV-FACTYDASKWATFWVPVIGILICAVGTFGDLCASMLK 287

QY 253 RDLGIGKMSNLLPHGGGLMDRLDGMPLPAAMVTWIL 288

Db 288 RDIGIKDMGHLLKGGGVMDRVDVSILMSAPFTCALL 323

RESULT 6

Q83G74 PRELIMINARY; PRT; 280 AA.

ID Q83G74 AC Q83G74;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Phosphatidate cytidyltransferase (EC 2.7.7.41).
 GN CDS A OR TW445.
 OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=203267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Raoult D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,
 RA Claverie J.-M.;
 RA "Tropheryma whipplei illustrates the diversity of gene loss patterns
 RT in small genome bacterial pathogens.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB016851; AA044542.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR000374; PC:trans.
 DR Pfam; PF01148; CTP_transf_1.
 DR PROSITE; PS01315; CDS; 1.
 KW Transferase; Nucleotidyltransferase; Complete proteome.
 SQ SEQUENCE 280 AA; 31290 MW; 6E64036636D5D6A9 CRC64;

 Query Match 20.6%; Score 322.5; DB 16; Length 280;
 Best Local Similarity 31.7%; Pred. No. 3.8e-17; Indels 37; Gaps 10;
 Matches 92; Conservative 47; Mismatches 114;
 QY 1 MNEPEQ-HRSMRMPKPNAGRDKAATAVGIGLVLLG--IVLSPWGYILVAGFM 57
 DB 1 MNEAMLRYPFMLKQVQARTGRDLVLAIGVGLLGLLISLVFLHLFVLVLCVGT 60
 QY 58 AAATWEGSRLEGGVHLEPLM-----IIGQAIWLSWPGTGMGLASFVATVLMY 112
 DB 61 CLA--EIFTATHTRGIFIVRPLTSTVLPVLAPFFASW---FLPSIAFFMLTLLAQF 115
 QY 113 FRIFYNGTEKARNYLRDTSVGIPLVTLPGSFAAMLSLMQNNISIPGTIFILTFMLCV 172
 DB 116 F-----LRARFSLH-----SLFLYIPITVSLFAIAAHT---GRLWVPFMTV 159
 QY 173 IASDVGYYAGVFFGSHPMAPIVSPKSWEGFAG----SIVLGSVTGALSVEFLDHHW 228
 DB 160 IASDTPSYVFGTLFGRHLLAPRISPNTKWEGLMGFFSFLFGTLTGILLHKSLL----L 215
 QY 229 MGVLGALVVCATLGLDVLVESQFKDLGKDMNSLLPGHGLMDRLDGM 278
 DB 216 FSATAASILFLFAILLGLAELSYIKRLGVKDMGSLPGHGLMDRVDNML 265

 RESULT 7
 ID Q83NN2 PRELIMINARY; PRT; 269 AA.
 AC Q83NN2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative integral membrane phospholipid biosynthetic
 DE nucleotidyltransferase.
 GN TW322.
 OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=218496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bentley S.D., Mairwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
 RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,
 RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
 RA Barrell B.G., Parkhill J., Rellman D.A.;
 RA

RT "Sequencing and analysis of the genome of the Whipple's disease
 RT bacterium Tropheryma whipplei";
 RL Lancet 361:637-644 (2003).
 DR EMBL; BX251411; CAD66995.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR000374; PC:trans.
 DR Pfam; PF01148; CTP_transf_1.
 DR PROSITE; PS01315; CDS; 1.
 KW Complete proteome.
 SQ SEQUENCE 269 AA; 23898 MW; B91FB3AC0FC604C6 CRC64;

 Query Match 20.6%; Score 322; DB 16; Length 269;
 Best Local Similarity 32.7%; Pred. No. 4e-17;
 Matches 88; Conservative 44; Mismatches 101; Indels 36; Gaps 9;
 QY 21 GRDLKAAIAVIGIGLVLLG--IVLSPWGYILVAGFAAATWEGSRLEGGVHPLP 78
 DB 11 GRDLVLAIGVGLLGLLISLVFLHLFVLVLCVLCVLCVLCVLCVLCVLCVLCV 68
 QY 79 IM-----IIGQAIWLSWPGTGMGLASFVATVLMYFRIFYNGTEKARNYLRDTSV 133
 DB 69 FLTLSTVLPVLAPFFASW---FLPSIAFFMLTLLAQF-----LRARFSLH---- 113
 QY 134 GIFVLTWIPFLGSAAMLSLMQNNISIPGTIFILTFMLCVIASDVGGYIAGVFFGSHMAP 193
 DB 114 SLFLYIPITVSLFAIAAHT---GRLWVPFMTVIASTFYSVFGTLFGRHLLAP 169
 QY 194 LVSPKSWEGFAG----SIVLGSVTGALSVEFLDHHWGMVLGALVVCATLGLDVL 249
 DB 170 RISFNKTWELGCGFFSFLFGTLTGILLHKSLL-----LFSATAASILFLFAILLGL 225
 QY 250 QFKRDLGKDMNSLLPGHGLMDRLDGM 278
 DB 226 YIKRLGVKDMGSLPGHGLMDRVDNML 254

 RESULT 8
 ID Q886N8 PRELIMINARY; PRT; 271 AA.
 AC Q886N8;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Phosphatidate cytidyltransferase.
 GN CDS A-1 OR PSPT01539.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
 RA White O., Fraser C., Collier A.;
 RA "Complete sequence of Pseudomonas syringae";
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AE016861; AAC55059.1; -
 DR TIGR; PSPT01539; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR000374; PC:trans.
 DR Pfam; PF01148; CTP_transf_1.
 DR PROSITE; PS01315; CDS; 1.
 KW Transferase; Nucleotidyltransferase; Complete proteome.
 SQ SEQUENCE 271 AA; 23308 MW; D737D0386868686B CRC64;
 KW

Db	219	YIIFVCFVCA-ILFSVIGDITESTFKESGKDTGNLIPGHGILDRMDSLSAIPFVT	277
Qy	287	ILSVI 291	
		:	
Db	278	LLFL 282	
RESULT 10			
Q86MH5		PRELIMINARY;	PRT; 271 AA.
ID	Q86MH5		
AC	Q86MH5;		
DT	01-JUN-2003 (TReMBLrel. 24, Created)		
DT	01-JUN-2003 (TReMBLrel. 24, Last sequence update)		
DT	01-OCT-2003 (TReMBLrel. 25, Last annotation update)		
DE	Phosphatidate cytidylyltransferase.		
GN	CDSA OR PPI596		
OS	Pseudomonas putida (strain Kt2440).		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;		
OC	Pseudomonadaceae; Pseudomonas.		
OX	NCBI_TaxID=160488;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22423060; PubMed=12534463;		
RA	Nelson K.E., Weinel C., Fouts I.T., Dodson R.J., Hilbert H.,		
RA	Martins dos Santos V.A.P., Paulsen D.E., Gill S.R., Pop M., Holmes M.,		
RA	Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,		
RA	Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,		
RA	Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,		
RA	Unterback T., Razzo M., Lee K., Kosack D., Moesli D., Wedler H.,		
RA	Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,		
RA	Kiewitz C., Eisen J., Timmis K.N., Duesterhoft A., Tuemmler B.,		
RA	Fraser C.M.;		
RA	"Complete genome sequence and comparative analysis of the		
RT	metabolically versatile Pseudomonas putida KT2440.";		
RL	Environ. Microbiol. 4:799-808(2002).		
DR	EMBL; AE016779; AAN67217.1; -.		
DR	TIGR; FPI596; -.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0004605; P:phosphatidate cytidylyltransferase activity; IEA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0008654; P:phospholipid biosynthesis; IEA.		
DR	InterPro; IPR000374; PC trans.		
DR	Pfam; PF01148; CTP transf. 1.		
DR	PROSITE; PS01315; CDS; 1.		
KW	Nucleotidyltransferase; Transferase; Complete proteome.		
SQ	SEQUENCE 271 AA; 29075 MW; 68DD533670B5EB3E CRC64;		
Query Match 19.1%; Score 299; DB 16; Length 271;			
Best Local Similarity 34.0%; Pred. No. 2.4e-15;			
Matches 100; Conservative 39; Mismatches 91; Indels 64; Gaps 15;			

```

22 RDLKAAIANGIGIGVLLLG-----IVLSPGWYIIVAGFMAAATWVSGRUK 70
   QY
5 RIITALLPVALGGFFLLNGGDFALFIGVVTLGAWEARL-AGLMAQPLRIAYAAVVA 63
   DB
71 GG----YHLPUPIMIIGQQAIILWLSWPFPGTMGILASFVATVLVLMYFRIFYNGTKEARN 126
   QY
64 GALLMLHLPELAPWVLGAIVN-WG-----LATLVLVTPR-----SSDLWASA 107
   DB
127 VLRDTSVGIFVLTIWLPFGSFAAMLISIMONNSIPGTIFYILTFMLCVIASDVGGYIAGVFF 186
   QY
108 ACR-LLIGLLVL--LPAWQG----LVLLXHWPL-GNWLISVWVLVWAADIGAYFSGRAF 159
   DB
187 GSHPMABLVPSPKSWEGFAGSIVLGSVTQALSVHFL-----LDHHWWMGVI-----LCA 236
   QY
160 GKRKLAPQVSPKSWEG-----VYGGGLAVSLIITLGVGISDDWGFQGLLGLLGA 210
   DB
237 LVVCAT-LGDLVPSQFKRDLIGIKDMNLNLIPIGHGGLMDRLDGLMLPA-----AMVTW 285
   QY
211 LLVMSVVVGLDTESMFRRSGIKDLSNLLPCHGGVLDRLDLSLTAIPIFAVLW 264
   DB

```


Result No.	Score	Query Match	Length	DB	ID	Description
1	298	19.1	500	4	US-09-252-991A-21214	Sequence 21214, A
2	295	18.9	289	4	US-09-489-039A-12390	Sequence 12390, A
3	285	18.2	249	4	US-09-282-218A-17	Sequence 17, Appl
4	282.5	18.1	292	4	US-09-543-681A-6495	Sequence 6495, Ap
5	262	16.8	285	4	US-09-328-353-7257	Sequence 7257, Ap
6	255	16.3	315	4	US-09-543-681A-7121	Sequence 7121, Ap
7	248	15.9	296	4	US-09-240-236-2118	Sequence 2118, Ap
8	247.5	15.8	261	4	US-09-134-001C-4966	Sequence 4966, Ap
9	246.5	15.8	346	4	US-09-252-991A-31107	Sequence 31107, A
10	219	14.0	192	4	US-09-134-000C-3752	Sequence 3752, Ap
11	199	12.7	446	2	US-08-672-814D-11	Sequence 11, Appl
12	199	12.7	446	3	US-09-333-696-11	Sequence 11, Appl
13	199	12.7	446	4	US-09-282-218A-19	Sequence 19, Appl
14	198.5	12.7	445	4	US-09-282-218A-12	Sequence 12, Appl
15	198.5	12.7	451	2	US-08-672-814D-2	Sequence 2, Appl
16	198.5	12.7	461	3	US-09-333-696-2	Sequence 2, Appl
17	195.5	12.5	461	4	US-09-282-218A-2	Sequence 2, Appl
18	193.5	12.4	308	4	US-09-198-452A-606	Sequence 606, App
19	142	9.1	457	4	US-09-282-218A-18	Sequence 18, Appl
20	116.5	7.5	238	4	US-09-107-532A-5419	Sequence 5419, Ap
21	115.5	7.4	326	3	US-08-986-768-2	Sequence 2, Appl
22	110	7.0	432	4	US-09-489-039A-9343	Sequence 8343, Ap
23	109.5	7.0	337	4	US-09-543-681A-5521	Sequence 5521, Ap
24	109.5	7.0	494	4	US-09-489-039A-11852	Sequence 11852, A
25	105.5	6.7	537	4	US-09-252-991A-24252	Sequence 24252, A
26	105.5	6.7	637	4	US-09-489-039A-10817	Sequence 10817, A
27	105	6.7	354	1	US-08-313-533-9	Sequence 9, Appl

Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 12390
 LENGTH: 289
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-12390

Query Match 18.9%; Score 295; DB 4; Length 289;
 Best Local Similarity 29.3%; Pred. No. 9.6e-23;
 Matches 85; Conservative 48; Mismatches 95; Indels 62; Gaps 11;
 30 VGGIGLVLLGVLSPMG-----WYILVAGFMAAAATWEGSRKEGGYHLPL 77
 30 MGFAIVLVVCLAAWENGQSGFTSTQRYWLVLCGLLAAMLFL---LPEYHYDVHQ 86
 78 PMITGGQAIWLSWPPFGTMGILASVATVLMYFRIFYNGTEKEARNYLRDTSV--GI 135
 87 PMV-----EGSLWASFAWM-----IVALLLVLSY-----PASAFAWRHRSKVLRLI 126
 136 F-VLTWIPLF-GSFAAMLSLQNNISIPGTYFTILFMLCVIASDVGGYIAGVFFGSHPMAP 193
 127 FGILITVFFGMLRALRAHWADNHYSALWLLYVWILVWGDADSGAYMFGKFKKLAP 186
 194 LVSPKSWEGFAGSIVLGSVGTGALSVEHFLDHHWVG-----VILGALV--VCA 241
 187 KVSFGKTVQGGFGLLTAAVIS-----WAGYVWHLDTPTVLLVCSVVAALAS 235
 242 TLGDLVESQFKEDIGIKMSNLLPGHGLMDRLDGLMPLPAAMVTWILSVI 291
 236 VLGDITSMFKREAGIKDGHLPFGHGLDRIDSLTAAPVFACLLLIV 285

RESULT 3

US-09-282-218A-17
 Sequence 17, Application US/09282218A
 Patent No. 6503700

GENERAL INFORMATION:
 APPLICANT: LEUNG, David W.
 TITLE OF INVENTION: MAMMALIAN CDP-DIACYLGLYCEROL SYNTHASE
 FILE REFERENCE: 077319/0153
 CURRENT APPLICATION NUMBER: US/09/282,218A
 CURRENT FILING DATE: 1999-03-31
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: Patent in ver. 2.0
 SEQ ID NO 17
 LENGTH: 249
 TYPE: PRT
 ORGANISM: E. coli
 US-09-282-218A-17

Query Match 18.2%; Score 285; DB 4; Length 249;
 Best Local Similarity 31.0%; Pred. No. 8.7e-22;
 Matches 85; Conservative 38; Mismatches 87; Indels 64; Gaps 12;
 47 WG-----WVILVAGFMAAAATWEGSRKEGGYHLPLPIMITGGQAIWLS--- 91
 7 WQQLSGFTTRSORVWLVLCGLLLALMLFL---LPEYHNTHQPELVIS-----LWASLW 59
 92 WFGTMGILASVATVLMYFRIFYNGTEKEARNYLRDTSVGI-FVLTWIPLF-GSFAA 149
 60 W-----IVALLLV-----FYPGSAAIWRN--SKTLRLIFGVLTIVPFFWGMAL 102
 150 MSLMQNNISIPGTYFTILFMLCVIASDVGGYIAGVFFGSHPMAPLVSPKSWEGFAGSIV 209

Db 103 RAWHYDENHYSALWLLYVWILVWGDADSGAYMFGKLFGRHKLAPKVPCKTWQGFIGGLA 162
 QY 210 LGSVTGALSVEHFLDHHWVG-----VILGALV--VCATLGDLVESQFKRDLCI 257
 Db 163 TAAVIS-----WYGMWANDVAPVILICSVIAALASVLDLTSMFKREAGI 211
 QY 258 KMSNLLPGHGLMDRLDGLMPLPAAMVTWILSVI 291
 Db 212 KDSCHLIPGHGILDRIDSLTAAPVFACLLLIV 245

RESULT 4

US-09-543-681A-6495
 Sequence 6495, Application US/09543681A
 Patent No. 6605709

GENERAL INFORMATION:
 APPLICANT: GARY BRETON
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709.1002-001
 CURRENT APPLICATION NUMBER: US/09/543,681A
 CURRENT FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: US 60/128,706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO 6495
 LENGTH: 292
 TYPE: PRT
 ORGANISM: Proteus mirabilis
 US-09-543-681A-6495

Query Match 18.1%; Score 282.5; DB 4; Length 292;
 Best Local Similarity 29.0%; Pred. No. 2e-21;
 Matches 88; Conservative 46; Mismatches 118; Indels 51; Gaps 10;

QY 22 RDLKAAIAYGIGLVLLGVLSPMGWYILVAGFMAAAATW-----VGSRLKEG- 71
 Db 10 RVITALLIPVIALFL---LPPAGFGLVIAISGLAGWMAQFIGHWSQKRIATGV 65
 QY 72 GYHLPLPIM-----ITGQAIWLSWPPFGTMGILASVATVLMYFRIFY 117
 Db 66 GFAALLVLMQASLPDPDHLDTAMIKNSLWAGLFW-----WVAAILVWSYP--- 112
 QY 118 NGTEKEARNYLRDTSVGI-FVLTWIPLF-GSFAAMLSLQNNISIPGTYFTILFMLCVIASD 176
 Db 113 NSANMKNVLLKLLPLG--LTIIVFYCGWALRMLGYSTNSYTGAWWLLYVWLLVWAD 170
 QY 177 VGGYIAGVFFGSHPMAPLVSPKSWEGFAGSIVLGSVGTGALSVEHFL-----LDHWWMGVI 232
 Db 171 SGAYAFGLRMKGHKMAPKVPCKTLEGLVGLITAGVWSWLFTRFAPITEVENHL---LL 227
 QY 233 LGCALVVCATLGDLVESQFKRDIGIKMSNLLPGHGLMDRLDGLMPLPAAMVTWILSVIS 292
 Db 228 ISGIVIVSVFGDLAESMPKRSVGINDSSQLIPGHGVLDRIDSLTAAPVFAGLVFLVS 287
 QY 293 SSY 295
 Db 288 SGF 290

RESULT 5

US-09-328-352-7257
 Sequence 7257, Application US/09328352
 Patent No. 6562958

GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252


```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5752
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5752

Query Match 14.0%; Score 219; DB 4; Length 192;
Best Local Similarity 40.0%; Pred. No. 4.9e-15;
Matches 48; Conservative 20; Mismatches 46; Indels 6; Gaps 2;

QY 165 ILTEMLCVI-ASDVGVIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHF-- 221
Db 61 VLLPFLFIWADIDGAYLFGRRFGRHKLMPDVPNKTIEGALGGILSAVVVAALFLVETA 120
QY 222 ---LDDHWMGVILGCALVVCATLGDVLSQFQKRDGKDKMSNLLPGHGLMDRLDGM 278
Db 121 NKGLFPYPMVPMVLTVLPSIVGQFGLVSESIKRHYGVKDSGNILPGHGLDRFDSLL 180

RESULT 11
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Weeks, Reitha
; TITLE OF INVENTION: MAMMALIAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 201 Elliott Avenue West
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44mb, double side, high density
; COMPUTER: PC Clone (486 microprocessor)
; OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT
; SOFTWARE: WORD 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,814D
; FILING DATE: 28-Jun-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Faciszewski, Stephen
; REGISTRATION NUMBER: 36,131
; REFERENCE/DOCKET NUMBER: 1803
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)282-7100
; TELEFAX: (206)284-6206
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Drosophila
; STRAIN:
; INDIVIDUAL ISOLATE:

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31107
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-31107

Query Match 15.8%; Score 246.5; DB 4; Length 346;
Best Local Similarity 25.3%; Pred. No. 1.5e-17;
Matches 80; Conservative 56; Mismatches 89; Indels 91; Gaps 16;

QY 33 GLGLVLLV---GIVL-----SPMGWYILVAGFMAATWVEGS--- 66
Db 46 GIGALLLASLVGFVVKWAGGPNVINDLNARINAWWVLTGP---AFW-LGQAAV 101
QY 67 -----RLKE-----GGYHLPDIPIMIGQAIWLSWPFETWGL 100
Db 102 ILLFYGVSFYALREFITLPTRRSDYPALVAAFVFLPMQYL---LIYDW----YSMF 153
QY 101 ASFVAT-VLVLMYFRIFYNGTEKEARNYLRDTS-----VGIFVLTWIPLFGSFAAMLS 152
Db 154 SIFIPYVFLLLPILASLGG---DTKHFLERASKVQWGLMLAVFCISFVP-----ALLT 204
QY 153 LMQNSNIPG-----TYFILTMLCVI-ASDVGVIAGVFFGSHPMAPLVSPKKSWEFGAGS 207
Db 205 L-----DIPGFEGRNLLIAYIVVQLSDVMQYICGLFGKRKIAPNLSPSKTVGEFVGG 260
QY 208 IVLGSVTGALSVEHFILDDHWMGVILGCALVVCATLGDVLSQFQKRDGKDKMSNLLPGH 267
Db 261 IALATAIGA-SLWNTTPNPLWQALIALIINLLFGGLVWSAIKRGRGVKDWGWHMIEGH 319
QY 268 GGLMDRLDGLMPLAAMV 283
Db 320 GGMLDRLDVSCFAAPI 335

RESULT 10
US-09-134-000C-5752
; Sequence 5752, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5752
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5752

Query Match 14.0%; Score 219; DB 4; Length 192;
Best Local Similarity 40.0%; Pred. No. 4.9e-15;
Matches 48; Conservative 20; Mismatches 46; Indels 6; Gaps 2;

QY 165 ILTEMLCVI-ASDVGVIAGVFFGSHPMAPLVSPKKSWEFGAGSIVLGSVTGALSVEHF-- 221
Db 61 VLLPFLFIWADIDGAYLFGRRFGRHKLMPDVPNKTIEGALGGILSAVVVAALFLVETA 120
QY 222 ---LDDHWMGVILGCALVVCATLGDVLSQFQKRDGKDKMSNLLPGHGLMDRLDGM 278
Db 121 NKGLFPYPMVPMVLTVLPSIVGQFGLVSESIKRHYGVKDSGNILPGHGLDRFDSLL 180

RESULT 11
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Weeks, Reitha
; TITLE OF INVENTION: MAMMALIAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 201 Elliott Avenue West
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44mb, double side, high density
; COMPUTER: PC Clone (486 microprocessor)
; OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT
; SOFTWARE: WORD 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,814D
; FILING DATE: 28-Jun-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Faciszewski, Stephen
; REGISTRATION NUMBER: 36,131
; REFERENCE/DOCKET NUMBER: 1803
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)282-7100
; TELEFAX: (206)284-6206
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Drosophila
; STRAIN:
; INDIVIDUAL ISOLATE:
```

DEVELOPMENTAL STAGE:

HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:

NAME/KEY: CDP-diacylglycerol synthase (CDS)

LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:

US-08-672-814D-11

Query Match 12.7%; Score 199; DB 2; Length 446;
Best Local Similarity 23.3%; Pred. No. 1.9e-12;
Matches 80; Conservative 40; Mismatches 111; Indels 112; Gaps 13;

QY 50 YILVAGMAAATWVGSRLEGGYHLPIMI-----IGGQ-----AIIW 89
Db 91 WIMICG-----ALIIYG-----PLAMITLLVQVKCFQELISIGYQVYRIHGLPW 138
QY 90 ---LSWPFGTMGILAS--EVATVILMYFRIFYNGTE--KEARNYLRTDTSVGIFVL--TW 140
Db 139 FRSLSWYF---LLTSNYFFYGENVLDYFGVIVNRVEYLKFLVYTHRFSLFALYIIGFVW 194
QY 141 IPL-----FGSFA-----AMLSIMQNNISPGTYFILTFLMCLVIASDVGGY 180
Db 195 FVLSLVKYYIKQSLFANWTHSVLLIVTQSYLLIQNIFEGLIWFIVPVSIMVNDWMAY 254
QY 181 IAGVFFGSHPMAPLVSPKKSWEAGFAGSIVLGSVTGALSVMHFLLDHWH----- 227
Db 255 VFGFFGRTPLIKL-SPKKTWEGFIGGFATVLFGLFSLVYLCNYQYFICPIQYSEQR 313
QY 228 -----WGVILGCVLVVCAITLGDLVESQFKRD 254
Db 314 MTMSCVPSYLPFTQEYSKLFGIGKTLNLYPFIWHISLSLSFSSIIIGFPGFFASGFKRA 373
QY 255 LGIKDMSNLLPGHGLMDRLDMLPAAMVTWLLISVSSSYPS 297
Db 374 FKIKDFGDMIPGHGIMDRFCQFLMATFNVNVIISFIRTPSPA 416

RESULT 12

US-09-333-696-11
Sequence 11, Application US/09333696
Patent No. 6200769

GENERAL INFORMATION:

APPLICANT: Leung, David W.
APPLICANT: Weeks, Reitha
TITLE OF INVENTION: MAMMALIAN
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cell Therapeutics, Inc.
STREET: 201 Elliott Avenue West
City: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98119

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" disk, 1.44Mb, double side, high density
COMPUTER: PC Clone (486 microprocessor)
OPERATING SYSTEM: MS-DOS version 6.1, Windows NT
SOFTWARE: WORD 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/333,696

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/672,814

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Paciszewski, Stephen
REGISTRATION NUMBER: 36,131
REFERENCE/DOCKET NUMBER: 1803

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)282-7100
TELEFAX: (206)284-6206
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 446

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: Drosophila
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:

NAME/KEY: CDP-diacylglycerol synthase (CDS)

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-333-696-11

Query Match 12.7%; Score 199; DB 3; Length 446;

Best Local Similarity 23.3%; Pred. No. 1.9e-12;

Matches 80; Conservative 40; Mismatches 111; Indels 112; Gaps 13;

QY 50 YILVAGMAAATWVGSRLEGGYHLPIMI-----IGGQ-----AIIW 89
Db 91 WIMICG-----ALIIYG-----PLAMITLLVQVKCFQELISIGYQVYRIHGLPW 138
QY 90 ---LSWPFGTMGILAS--EVATVILMYFRIFYNGTE--KEARNYLRTDTSVGIFVL--TW 140
Db 139 FRSLSWYF---LLTSNYFFYGENVLDYFGVIVNRVEYLKFLVYTHRFSLFALYIIGFVW 194
QY 141 IPL-----FGSFA-----AMLSIMQNNISPGTYFILTFLMCLVIASDVGGY 180
Db 195 FVLSLVKYYIKQSLFANWTHSVLLIVTQSYLLIQNIFEGLIWFIVPVSIMVNDWMAY 254
QY 181 IAGVFFGSHPMAPLVSPKKSWEAGFAGSIVLGSVTGALSVMHFLLDHWH----- 227
Db 255 VFGFFGRTPLIKL-SPKKTWEGFIGGFATVLFGLFSLVYLCNYQYFICPIQYSEQR 313
QY 228 -----WGVILGCVLVVCAITLGDLVESQFKRD 254
Db 314 MTMSCVPSYLPFTQEYSKLFGIGKTLNLYPFIWHISLSLSFSSIIIGFPGFFASGFKRA 373
QY 255 LGIKDMSNLLPGHGLMDRLDMLPAAMVTWLLISVSSSYPS 297
Db 374 FKIKDFGDMIPGHGIMDRFCQFLMATFNVNVIISFIRTPSPA 416

RESULT 13

US-09-282-218A-19

Sequence 19, Application US/09282218A

Patent No. 6503700

GENERAL INFORMATION:

APPLICANT: Leung, David W.
TITLE OF INVENTION: MAMMALIAN CDP-DIACYLGLYCEROL SYNTHASE
FILE REFERENCE: 077319/0153
CURRENT APPLICATION NUMBER: US/09/282,218A
CURRENT FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 19
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 19
LENGTH: 446

Qy 56 FMAAATWE---VGSRLKEGGYHLPLPIMIIGGQAIIM---LSWPFGTGMGILASEFVATVLV 109
Db 118 IOVKCFHEIITIGYRVYH-SYDLP-----WFRTLWYF-----LLC 152
Qy 110 LMYFRIFYNCT-----EKEAR-----NYLRDTSVGI-----FVLTWIP-----L 143
Db 153 VNYF--PYGETVADYFATFVOREEQQLRIRYHRFISFALYLAGPCMFVLSLVKEHYRLQ 210
Qy 144 FGSFA-----AMLSLQN-----NSIPGTYFILTFLMLCVIASDVGGYIAGVREGSHPMAP 193
Db 211 FYMFANWHTVLLITVTQSHLVIQNLFEGMWFLVPISSVICNDITAYLFGFFFGRTPLIK 270
Qy 194 LVSPKKSWEFGAGSIVLGSVTGALSVMHFLDDHHWM-----GVILGC----- 235
Db 271 L-SPKKTWEGFIGGFSTVVFGLAAVLSKYQYFVCPVEYRSDVNSFVTECPSELFOL 329
Qy 236 -----ALVVCATL-----GDLVESQFKRDLGDKMSNL 263
Db 330 QTYSLPPLXAVLRQERSLYPFQIHSLALSTFASLIGPFPGFASGFKRAFKIKDFANT 389
Qy 264 LPGHGGIMDRLD-CMLPAAWTVLILSVISSYPS 297
Db 390 IPGHGGIMDRDCQYLMATLVHGYITSFIRGNPS 424

Search completed: August 11, 2004, 13:36:57
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 15:46:53 ; Search time 8196 Seconds
(without alignments)
6874.810 Million cell updates/sec

Title: US-09-853-641-1
Perfect score: 1300
Sequence: 1 gaagtggttcgcgaagga.....aggaaaccgattccggctgg 1300

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1300	100.0	1300	6	AX297999	Sequence
C 2	1300	100.0	337200	1	AP005280	Coryneb
C 3	1300	100.0	349980	6	AX127149	Sequence
C 4	1300	100.0	349980	6	AX127150	Sequence
C 5	1193	91.8	1194	1	CX1312009	Coryneb
C 6	978	75.2	978	6	AX064835	Sequence
C 7	978	75.2	978	6	AX066517	Sequence
C 8	855	65.8	855	6	AX122300	Sequence
C 9	855	65.8	855	6	BD164417	Novel pol
C 10	596.8	45.9	303250	1	AP005220	Coryneb
C 11	342.2	26.3	348408	1	BX248358	Coryneb
C 12	191.8	14.8	16952	1	AE007118	Corybacte
C 13	191.8	14.8	318050	1	BX248344	Mycobacte
C 14	191.8	14.8	348676	15	BX842581	Bx248344 Mycobacteri
C 15	164	12.6	36985	1	MSGB1529CS	L78824 Mycobacteri
C 16	164	12.6	318200	1	MLSPRTN6	AL583922 Mycobacte
C 17	123.4	9.5	297	6	AX307920	Sequence
C 18	117	9.0	258	6	AX122299	Sequence
C 19	117	9.0	258	6	BD164416	Novel pol
C 20	110.2	8.5	299550	1	AP005031	Streptomy
C 21	107	8.2	308050	1	SC0939124	Streptomy
C 22	75.6	5.8	349981	1	BX572602	Rhodopseu
C 23	73.8	5.7	10029	1	AE012236	Xanthomon
C 24	72.4	5.6	300880	1	AE016917	Chromobac
C 25	72	5.5	299050	1	BX251411	Tropherym
C 26	72	5.5	302529	1	AE016851	Tropherym
C 27	71.8	5.5	10242	1	AE011773	Xanthomon
C 28	69.8	5.4	10771	1	AE014782	Bifidobac
C 29	69.8	5.4	349980	6	AX492787	Sequence
C 30	69.8	5.4	349980	6	AX553954	Sequence
C 31	69.6	5.4	300600	1	AP005352	Bradyrhiz
C 32	68.4	5.3	310029	1	AE016861	Pseudomon
C 33	67.2	5.2	301995	1	AE016779	Pseudomon
C 34	65.8	5.1	1845	1	PSECDFDS	Pseudomonas
C 35	65.8	5.1	14537	1	AE004785	Pseudomon
C 36	65	5.0	11455	1	AE012802	Chlorobiu
C 37	64.4	5.0	295250	1	BX321862	Nitrosomo
C 38	64	4.9	342805	1	BX571861	Photorhab
C 39	64	4.9	349980	6	AX770908	Sequence
C 40	63.8	4.9	300817	1	AE016943	Bacteroid
C 41	62.8	4.8	299350	1	AP006572	Gloeobact
C 42	60.8	4.7	10487	1	AE003942	Xylella f
C 43	60.8	4.7	300029	1	AE012554	Xylella f
C 44	59	4.5	10301	1	AE004297	Vibrioc
C 45	58.8	4.5	11002	1	AE004681	Pseudomon

ALIGNMENTS

RESULT 1
AX297999
LOCUS AX297999
DEFINITION Sequence 1 from Patent WO0183785.
ACCESSION AX297999
VERSION AX297999.1 GI:17128085
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nampoothiri, M., Moeckel, B., Pfeifferle, W., Eggeling, L. and Sahm, H.
TITLE Nucleotide sequences coding for the cda gene
JOURNAL Patent: WO 0183785-A 1 08-NOV-2001

linear PAT 26-NOV-2001

Degussa AG (DB) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)

FEATURES

source

Location/Qualifiers

1..1300

/organism="Corynebacterium glutamicum"

/mol_type="unassigned DNA"

/db_xref="taxon:1718"

187..195

RBS

gene

/gene="cdsA"

200..1093

CDS

/gene="cdsA"

/codon_start=1

/transl_table=11

/protein_id="CAD12679.1"

/db_xref="GI:17128086"

/translation="MNEPEQHRMRMPKPNKAGRLKAAIAVIGLVLIGIVL

SPWGYILVAGFMAAAWEVGRLEKGYHLPLPIMLIGGOAIIWLSPFGTWMGLAS

FVATVLMVFRIFVNDGEKARNYLRDTSVGI FVLTWLPFGSFAAMLILMONNSIP

GTVELFMLECVIADVGGYIAGVFFGSHPKAPLVPKSKSEFPAGSIVLGSVMTGLS

VHFLDHRHWMGVILGALVVCATILGDLVFSQFRDLGIKIDMSNLLPFGHGLMDRLDG

MLPAAVWTLILSVISSYPS"

ORIGIN

Query Match 100.0%; Score 1300; DB 6; Length 1300;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTCGTGTCGCAAGAAAAGAACTGATGAGGCTTAGAAGACTTTATCGCAATG 60
 DB 1 GAAGTCGTGTCGCAAGAAAAGAACTGATGAGGCTTAGAAGACTTTATCGCAATG 60

QY 61 GCTTAAGAAATACGTCGTTCAAGCTGATGGCGGGGAAACACACGCTTCTTTTCGCT 120
 DB 61 GCTTAAGAAATACGTCGTTCAAGCTGATGGCGGGGAAACACACGCTTCTTTTCGCT 120

QY 121 GCAAGAGTGTGGGAAGAAATTTTTCGAAAATGCTGGCACCATCAACAGTGACATTGTA 180
 DB 121 GCAAGAGTGTGGGAAGAAATTTTTCGAAAATGCTGGCACCATCAACAGTGACATTGTA 180

QY 181 GAACTTCAAGAGAACCCATGATGAACCGAGCAACATCACCGGTCCATGAGATGCC 240
 DB 181 GAACTTCAAGAGAACCCATGATGAACCGAGCAACATCACCGGTCCATGAGATGCC 240

QY 241 CAACCCAAAATAATGCGGTCGAGATCTCAAAGCTGCCATTCCTGTGGGATCGGACT 300
 DB 241 CAACCCAAAATAATGCGGTCGAGATCTCAAAGCTGCCATTCCTGTGGGATCGGACT 300

QY 301 GGGGTCCTCGTTCTTTTGGGGATGTCCTAAGCCCATGGGGTTGGTACATCCTCGTTGC 360
 DB 301 GGGGTCCTCGTTCTTTTGGGGATGTCCTAAGCCCATGGGGTTGGTACATCCTCGTTGC 360

QY 361 AGTTTATGCTGCAGCAACATCGGAAGTTGGTAGAGACTTAAAGAACGGCGCTATCA 420
 DB 361 AGTTTATGCTGCAGCAACATCGGAAGTTGGTAGAGACTTAAAGAACGGCGCTATCA 420

QY 421 TTTGCCACTGCCGATATGATCATCGCGGTCAGGCAATCATCTGGCTGTCATGGCCATT 480
 DB 421 TTTGCCACTGCCGATATGATCATCGCGGTCAGGCAATCATCTGGCTGTCATGGCCATT 480

QY 481 TGGCAGATGGCAATTTTGGCGTCTTTTGGCCACTGTGTGGTGTGATGATTTCCG 540
 DB 481 TGGCAGATGGCAATTTTGGCGTCTTTTGGCCACTGTGTGGTGTGATGATTTCCG 540

QY 541 AATTTTCTAATGCGACGAAAAGAACCGCAACTATTATGAGGACACCTCTGTGGG 600
 DB 541 AATTTTCTAATGCGACGAAAAGAACCGCAACTATTATGAGGACACCTCTGTGGG 600

QY 601 CATCTTCGTGTCACCTGGATTCATTTTGGAGCTTCGCTCGGATGTGCTGCTGAT 660
 DB 601 CATCTTCGTGTCACCTGGATTCATTTTGGAGCTTCGCTCGGATGTGCTGCTGAT 660

QY 661 GCAAAACAATCCATCCCGGGTACATATTTCATTTTGACGTTTCATGCTGTGTGATCGC 720

DB 661 GCAAAACAATCCATCCCGGGTACATATTTCATTTTGACGTTTCATGCTGTGTGATCGC 720
 QY 721 ATCGAGTGTGGCGGGTATATCGCGGGTGTGTCTTTTGGATCGCACCAATGGCCCGTT 780
 DB 721 ATCGAGTGTGGCGGGTATATCGCGGGTGTGTCTTTTGGATCGCACCAATGGCCCGTT 780
 QY 781 GGTAGTCCGAAGAAGTCTTTGGGAAGGCTTTCGGGCTCCATTGTCTTAGGATCGGTAC 840
 DB 781 GGTAGTCCGAAGAAGTCTTTGGGAAGGCTTTCGGGCTCCATTGTCTTAGGATCGGTAC 840
 QY 841 TGGTGCACTCAGTGTTCATCTCTGCTCGATCACCATCGGTGTGGTGTGATCTTGGG 900
 DB 841 TGGTGCACTCAGTGTTCATCTCTGCTCGATCACCATCGGTGTGGTGTGATCTTGGG 900
 QY 901 TTGTGCCCTAGTGTGTGGCGCACCGTGGTGACCTTTGAGTTCGAGTTCGAGTTCGAGTTC 960
 DB 901 TTGTGCCCTAGTGTGTGGCGCACCGTGGTGACCTTTGAGTTCGAGTTCGAGTTCGAGTTC 960
 QY 961 TTTGGGCATCAAGGATATGTCGAACTTCTTCAGGCCACGGCGGATTCGATGACCGTTT 1020
 DB 961 TTTGGGCATCAAGGATATGTCGAACTTCTTCAGGCCACGGCGGATTCGATGACCGTTT 1020
 QY 1021 GGATGTCATGCTCCCGCGCGGATGGTGACGTTGATCTCTGAGTGTGATCAGCAGCTC 1080
 DB 1021 GGATGTCATGCTCCCGCGCGGATGGTGACGTTGATCTCTGAGTGTGATCAGCAGCTC 1080
 QY 1081 GTATCGCTGTAAGCTTTGGGCGAGCTTTAAGTTCAAAAGCTTTGAAAGGCGCTGAGTG 1140
 DB 1081 GTATCGCTGTAAGCTTTGGGCGAGCTTTAAGTTCAAAAGCTTTGAAAGGCGCTGAGTG 1140
 QY 1141 CATAAGCTGTGACCTCAGCGCTTTTGGCTGTCAAAGTCTTAAAGGCTTTACGGAATT 1200
 DB 1141 CATAAGCTGTGACCTCAGCGCTTTTGGCTGTCAAAGTCTTAAAGGCTTTACGGAATT 1200
 QY 1201 TTCTTAACCTGGCAGGCTACTCAACATCGCGCACCCCAACAGCGCCATATCAATG 1260
 DB 1201 TTCTTAACCTGGCAGGCTACTCAACATCGCGCACCCCAACAGCGCCATATCAATG 1260
 QY 1261 CACCGTAAATGCTGCTAGTAGGAACCGATTCGCGGTGG 1300
 DB 1261 CACCGTAAATGCTGCTAGTAGGAACCGATTCGCGGTGG 1300

RESULT 2

AP005280/c

LOCUS

DEFINITION

AP005280 337200 bp DNA linear BCT 08-AUG-2002

AP005280 BA000036

AP005280.2 GI:22138850

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..337200

/organism="Corynebacterium glutamicum ATCC 13032"

AP005280 337200 bp DNA linear BCT 08-AUG-2002

Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 7/10.

AP005280 BA000036

AP005280.2 GI:22138850

Corynebacterium glutamicum ATCC 13032

Corynebacterium glutamicum ATCC 13032

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacteriineae; Corynebacteriaceae; Corynebacterium.

1 Nakagawa,S.

Complete genomic sequence of Corynebacterium glutamicum ATCC 13032

2 (bases 1 to 337200)

Direct Submission

Nakagawa,S.

Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.

Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,

Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com,

Tel:81-44-829-3031, Fax:81-44-813-1651)

On Aug 8, 2002 this sequence version replaced gi:21324668.

This sequence is conducted by collaboration of Kyowa Hakko Kogyo

Co. Ltd. And Kitasato University.

Location/Qualifiers

1..337200

/organism="Corynebacterium glutamicum ATCC 13032"

/mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"
 complement (518..1216)
 /gene="Cg11899"
 complement (518..1216)
 /gene="Cg11899"
 /codon_start=1
 /transl_table=11
 /product="Uncharacterized ACR"
 /protein_id="BAB99292.1"
 /db_xref="GI:21324669"
 /translation="MSELDIQLKQLQRYSQWAVRAIPALGALDDRRTEVTDQAAKFFA
 DLEAGKTVIGTINASGRADADMIWHAAEFEDIQAFADFRITILQGVSEVFW
 IGNALHRESENKAHLPSFINGEBAKDWITVYFVRSYDVMTPSEPLKRSILREHGA
 AVFEPDVAIVTAPALGDYEWLAFAEDELHRLVDLMHKRYTEARLHVREELFFIS
 QQRVDIALIKVLP"
 1489..2115
 /gene="Cg11900"
 1489..2115
 /gene="Cg11900"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB99293.1"
 /db_xref="GI:21324670"
 /translation="MQIVIDSEATSOHKTSATPAESTPAPSEAVESMHRARLPBLT
 LGTRIPORLAPFHAIGLEVNGRESDSDVSTNEGDSFGRLILHDPGAETWEGAM
 RLVAIQADMHVASDPLLEVAWQMEGLEQAGAGFNLGTVTSTTSVRFGEIG
 GPSPAYQVEMRASVTATQTDLTAHVEAFAPAAVLASVGLPPEGVTELR"
 2072..3337
 /gene="Cg11901"
 2072..3337
 /gene="Cg11901"
 /EC_number="3.1.26.3"
 /note="PF00570:HRDC domain
 PF01612:3'-5' exonuclease"
 /codon_start=1
 /transl_table=11
 /product="Ribonuclease D"
 /protein_id="BAB99294.1"
 /db_xref="GI:21324671"
 /translation="MLDPQPARSPNYEGRLDITWVSDLLQPRDGIPLLSTPGFTAA
 DLLASGTGFFAIDTERASGFDYDRAFLQIRRRSGTLLFDPEQRPPLTQALKPVL
 NGQWIIHAASGTDLSLAWLDLHPLGLFDTELAGRLAGFDHVLAAWVEOIFDLHLK
 GHSEDSKRLPESNLAYALDVMLELADVAEILDDQGLPMAEQSFVHIDVF
 ATMPESKTMQDLKGLSTLKPQOLVAREMWERLERSPAASRDLAPCKVLSKNVIVE
 VARVLPTPABLAQVKGPPGSRQATKRFRITITALKSPRRNWFKPQQRQDGIPIRR
 AWASYIPBEHVQLEIRALIDDLADINVPGENILQPSLTRVAVWMAKHTGETHNAET
 LNAVIRYQASQWQIDQTFPILSANLLKL"
 complement (3399..5309)
 /gene="Cg11902"
 complement (3399..5309)
 /gene="Cg11902"
 /note="PF02779:Transketolase, pyridine binding domain
 PF02780:Transketolase, C-terminal domain
 TIGR00204:dxs: deoxyxylulose-5-phosphate synthase"
 /codon_start=1
 /transl_table=11
 /product="Deoxyxylulose-5-phosphate synthase"
 /protein_id="BAB99295.1"
 /db_xref="GI:21324672"
 /translation="MGILNISTPADIKALNDEDLALAKEIRTFLLVKVAATGGHGL
 PNLGVELTILHGVFSDPODPIFDTSQSYVHKILTGRAKDFDSLRQDGLSGYTC
 PNESEHMTSSHAALSADGLSKAQLDGDDTHSVAVAGDGLATGMCWEALNN
 IAAGRDKRVVNVNDGKRSYPTTIGFAENLAGLRMQPFYDREKGTSLKSMGWVG
 BRTFEALHAKFEGVKSTVITPEMPELGMKVGPVDCGNOKAVDNALKYADYDGLI
 VHMVTEKRGYAPAEQDLDELHSTGIVDPLTGPASAKPGTWSFSDSLVKIQAQ
 ENVAIVTAMAGPTGLSKFRANFNRFVDAIEQAHVATSAAGLALGCKHPVAVIST
 FLNAPQOLLMDVMKQPVTLVLDNRGVTGSGDASHGNYDMWALTSIVPGVQAAPR
 DEDSLRLNEAISIDGPTVTRPKGLDPTPIVAIDTLEDGVDVLAIVEDSTVST
 DAPSLVIAUGERATVADVASRIKQGVNVTVDDRWIVPIPOSILVALSDHDDHVTIT

/mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"
 complement (518..1216)
 /gene="Cg11899"
 complement (518..1216)
 /gene="Cg11899"
 /codon_start=1
 /transl_table=11
 /product="Uncharacterized ACR"
 /protein_id="BAB99292.1"
 /db_xref="GI:21324669"
 /translation="MSELDIQLNKLQRYSQWAVRAIPGALDDRRTEVTDQAAKFFA
 DLEAGKTVIGTINASGRADADMIWHAAEFEDIQAFADFRITILQGVSEVFW
 IGNALHRESENKAHLPSFINGEBAKDWITVYFVRSYDWTIMPSLKRSLIREHGA
 AVFEPDVAINTVPAPALGDYEWLAFAEDELHRLVDLMHKRYTEARLHVREELFFIS
 QQRVDIALIKVLP"
 1489..2115
 /gene="Cg11900"
 1489..2115
 /gene="Cg11900"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB99293.1"
 /db_xref="GI:21324670"
 /translation="MQIVIDSEATSOHKTSATPAESTPAPSEAVESMHRARLPBLT
 LGTRIPPORLAPFHAIGLEVNGRESDSDVNSTEGDSFGRLILHDPGAETWEGAM
 RLVAIQADMHVASDPLLEVAWQMEGLEQAGAGFNLGTVTSTTSVRFGEIG
 GPSPAYQVEMASATATQTDLTAHVEAFAPAAVLASVAGLPPEGVTELR"
 2072..3337
 /gene="Cg11901"
 2072..3337
 /gene="Cg11901"
 /EC_number="3.1.26.3"
 /note="PF00570:HRDC domain
 PF01612:3'-5' exonuclease"
 /codon_start=1
 /transl_table=11
 /product="Ribonuclease D"
 /protein_id="BAB99294.1"
 /db_xref="GI:21324671"
 /translation="MLDPQPARSPNYEGRLDITWVSDLLQPRDGIPLLSTPGFTAA
 DLLASGTGFFAIDTERASGFDYDRAFLQIRRRSGTLLFDPEQFPPLTQALKPVL
 NGQWIIHAASGTDLSLAWLDLHGLLFDTELAGRLAGFDHVLAAWVEOIFDLHLK
 GHSEDSKRLPESNLAYALDVMLELADVAEILDDQGLPMAEQSFVHVDQF
 ATMPESKTMQDLKGLSTLKPQOLVAREMWLERSPASRDLAPKVLNKNVIVE
 VARVLPTPABLAQVKGPPGQSQATKRFRITITALKSPRRNWFKQQRQDGIPIRR
 AWASYIPBEHVQLEIRALIDDLADINVPGENILQPSLTLRVAVMAKHTGETHNAET
 LNAVIRYQASQWQIDQTFPILSANLLKL"
 complement (3399..5309)
 /gene="Cg11902"
 complement (3399..5309)
 /gene="Cg11902"
 /note="PF02779:Transketolase, pyridine binding domain
 PF02780:Transketolase, C-terminal domain
 TIGR00204:dxs: deoxyxylulose-5-phosphate synthase"
 /codon_start=1
 /transl_table=11
 /product="Deoxyxylulose-5-phosphate synthase"
 /protein_id="BAB99295.1"
 /db_xref="GI:21324672"
 /translation="MGILNSTPADIKALNDEDLALAKEIRTFLLVKVAAATGGHLG
 PNLGVELTILHGVFSDPODPIFDTSQSYVHKILTGRAKDFDSLRQDGLSGYTC
 PNESEHMTSSHAALSYADGLSKAQLDGDDTHSVAVGDSALGTGMCWEALNN
 IAAGRDKRVVNVNDGKRSYPTTIGFAENLAGLRMQPFYDRFWEKGTSLKSMGVVG
 BRTFEALHAKFEGVKSTVITPEMPELGMKVGPVDCGNOKAVDNALKYADYDGLI
 VMVTEKRGYAPAEQDLDELHSTGTVIDPLTGPASAKPGTWSFSDBLVKIQAQ
 ENVAITRAMAGPTGLSKFRANFNRFVDGIAEQHAVISAAGLALGCKHPVAVIST
 FLNAFQOLLMDVMKQPVTLVLDNRGVTGSGDSAGHNVMDALTSIVPGVQAAPR
 DEDSLRLNEAISIDGPTVTRPKGLDPTPIVAIDTLEDGVDVLAEDVSTEDST
 DAPSVLIAGGERATVADVASRIKQGVNVTVDDRWIPEIPOSILVALSDHDDHVTIT

/mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"
 complement (518..1216)
 /gene="Cg11899"
 complement (518..1216)
 /gene="Cg11899"
 /codon_start=1
 /transl_table=11
 /product="Uncharacterized ACR"
 /protein_id="BAB99292.1"
 /db_xref="GI:21324669"
 /translation="MSELDIQLNKLQRYSQWAVRAIPGALDDRRTEVTDQAAKFFA
 DLEAGKTVIGTINASGRADADMIWHAAEFEDIQAFADFRITILQGVSEVFW
 IGNALHRESENKAHLPSFINGEBAKDWITVYFVRSYDWTIMPSLKRSLIREHGA
 AVFEPDVAINTVPAPALGDYEWLAFAEDELHRLVDLMHKRYTEARLHVREELFFIS
 QQRVDIALIKVLP"
 1489..2115
 /gene="Cg11900"
 1489..2115
 /gene="Cg11900"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB99293.1"
 /db_xref="GI:21324670"
 /translation="MQIVIDSEATSOHKTSATPAESTPAPSEAVESMHRARLPBLT
 LGTRIPPORLAPFHAIGLEVNGRESDSDVNSTEGDSFGRLILHDPGAETWEGAM
 RLVAIQADMHVASDPLLEVAWQMEGLEQAGAGFNLGTVTSTTSVRFGEIG
 GPSPAYQVEMASATATQTDLTAHVEAFAPAAVLASVAGLPPEGVTELR"
 2072..3337
 /gene="Cg11901"
 2072..3337
 /gene="Cg11901"
 /EC_number="3.1.26.3"
 /note="PF00570:HRDC domain
 PF01612:3'-5' exonuclease"
 /codon_start=1
 /transl_table=11
 /product="Ribonuclease D"
 /protein_id="BAB99294.1"
 /db_xref="GI:21324671"
 /translation="MLDPQPARSPNYEGRLDITWVSDLLQPRDGIPLLSTPGFTAA
 DLLASGTGFFAIDTERASGFDYDRAFLQIRRRSGTLLFDPEQFPPLTQALKPVL
 NGQWIIHAASGTDLSLAWLDLHGLLFDTELAGRLAGFDHVLAAWVEOIFDLHLK
 GHSEDSKRLPESNLAYALDVMLELADVAEILDDQGLPMAEQSFVHVDQF
 ATMPESKTMQDLKGLSTLKPQOLVAREMWLERSPASRDLAPKVLNKNVIVE
 VARVLPTPABLAQVKGPPGQSQATKRFRITITALKSPRRNWFKQQRQDGIPIRR
 AWASYIPBEHVQLEIRALIDDLADINVPGENILQPSLTLRVAVMAKHTGETHNAET
 LNAVIRYQASQWQIDQTFPILSANLLKL"
 complement (3399..5309)
 /gene="Cg11902"
 complement (3399..5309)
 /gene="Cg11902"
 /note="PF02779:Transketolase, pyridine binding domain
 PF02780:Transketolase, C-terminal domain
 TIGR00204:dxs: deoxyxylulose-5-phosphate synthase"
 /codon_start=1
 /transl_table=11
 /product="Deoxyxylulose-5-phosphate synthase"
 /protein_id="BAB99295.1"
 /db_xref="GI:21324672"
 /translation="MGILNISTPADIKALNDEDLALAKEIRTFLLVKVAATGGHLG
 PNLGVELTILHGVFSDPODPIFDTSQSYVHKILTGRAKDFDSLRQDGLSGYTC
 PNESEHMTSSHAALSYADGLSKAQLDGDDTHSVAVGDSALGTGMCWEALNN
 IAAKGRARVYVNVNDKRSYPTTIGFAENLAGLRMQPFYDREKGTSLKSMGWVG
 BRTTEALHAKFEGVKSTVITPEMPELGMKVGPVDCGNOKAVDNALKYADYDGLI
 VMVTEKRGYAPAEQDLDELHSTGTVIDPLTGPASAKPGTWSFSDBLVKIQAQ
 ENVAITARMAGPTGLSKFRANFNRFVDAIEQAHVATSAAGLALGGRKFPVAIVST
 FLNAPQOLLMDVMKQPVTLVLDNRSGVTSQSGASHGVDWALTSIVPGVQAAPR
 DEDSLRLNEAISIDGPTVYRPPKGLPTPIVAIDTLEDGVDVLAIVEDSTVST
 DAPSVLIAGGERATVADVASRIKQGVNVTVDDRWIVPIPOSILVALSDHDDHVTIT

/product="Hypothetical protein"
 /protein_id="BAB99300.1"
 /db_xref="GI:21324677"
 /translation="MNINHTYQVSWSEDDQEFATVAEPFSLSWLFSQDQAEABAGLR
 NLVAEVLDDMKQTGEQIPTSVGRNL"
 complement (9277..9570)
 /gene="Cg11908"
 CDS
 complement (9277..9570)
 /gene="Cg11908"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB99301.1"
 /db_xref="GI:21324678"
 /translation="MADYDPRRRVDELETSLEGKAVENANSMDDDGRIVRSF
 BIPNVLSGEELNVVPPRADEFTCASCLVQRNRKSHVEPDDGSIICLDC"
 complement (9721..10539)
 /gene="Cg11909"
 CDS
 complement (9721..10539)
 /gene="Cg11909"
 /note="PF00459: Inositol monophosphatase family"
 /codon_start=1
 /transl_table=11
 /product="Archaeal fructose-1,6-bisphosphatase and related
 enzymes of inositol monophosphatase family"
 /protein_id="BAB99302.1"
 /db_xref="GI:21324679"
 Query Match 100.0%; Score 1300; DB 1; Length 337200;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAGTCGTTGTCGCAAGAAAAGAACTGATGAGGTCCTAGAGACCTTTATCGCAATG 60
 DB 137324 GAAGTCGTTGTCGCAAGAAAAGAACTGATGAGGTCCTAGAGACCTTTATCGCAATG 137265
 QY 61 GCTTAAGAATTACGTCGTTCAACGTCGATGAGGCGGGGAAACGACGCTTCTTTTCGTT 120
 DB 137264 GCTTAAGAATTACGTCGTTCAACGTCGATGAGGCGGGGAAACGACGCTTCTTTTCGTT 137205
 QY 121 GCAAGAGTGTGTTGGAAGAAATTTTTCGAAATGCTGGCAACCATCAACAGTGACACTGTTA 180
 DB 137204 GCAAGAGTGTGTTGGAAGAAATTTTTCGAAATGCTGGCAACCATCAACAGTGACACTGTTA 137145
 QY 181 GAAATCTCAAGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCATGAGATGCC 240
 DB 137144 GAAATCTCAAGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCATGAGATGCC 137085
 QY 241 CAAACCCAAAATAATGCGGTGCGAGATCTCAAGCTGCCATTCGTTGGGATCGGACT 300
 DB 137084 CAAACCCAAAATAATGCGGTGCGAGATCTCAAGCTGCCATTCGTTGGGATCGGACT 137025
 QY 301 GCGCGTCCTCGTTCTTTTGGGATTTGCTTAAGCCCATGGGTTGGTACATCCTCGTTGC 360
 DB 137024 GCGCGTCCTCGTTCTTTTGGGATTTGCTTAAGCCCATGGGTTGGTACATCCTCGTTGC 136965
 QY 361 AGGTTTATGCGTCAGCAACATGGAAGATGTTGTTAGAGACTTAAGAGAGCGGCTATCA 420
 DB 136964 AGGTTTATGCGTCAGCAACATGGAAGATGTTGTTAGAGACTTAAGAGAGCGGCTATCA 136905
 QY 421 TTTGCACTGCCGATTATGATCATCGCGGTCAGGCAATCATCTGGCTGTATGCCCAT 480
 DB 136904 TTTGCACTGCCGATTATGATCATCGCGGTCAGGCAATCATCTGGCTGTATGCCCAT 136845
 QY 481 TGGCAGATGGGCATTTGGCGCTTTTGTGGCCTCTGTTGGTGTGATGATTTCCG 540
 DB 136844 TGGCAGATGGGCATTTGGCGCTTTTGTGGCCTCTGTTGGTGTGATGATTTCCG 136785
 QY 541 RAATTTCTACATGACCGGAAAAGAGCCCGCAACTATTGAGGACACCTCTGTGGG 600
 DB 136784 RAATTTCTACATGACCGGAAAAGAGCCCGCAACTATTGAGGACACCTCTGTGGG 136725
 QY 601 CATCTCGTGTCTCACCTGGATTCCATTGTTGGGAAGCTTCGTCGATGCTGTCTGAT 660

DB 136724 CATCTCGTCTCACCTGGATTCCATTGTTGGAAGCTTCGCTGATGCTGCTGAT 136665
 QY 661 GCAAAACAATTCATCCCGGTACATATTTTCAATTTTGAGCTTCATGCTGTGTGATCGC 720
 DB 136664 GCAAAACAATTCATCCCGGTACATATTTTCAATTTTGAGCTTCATGCTGTGTGATCGC 136605
 QY 721 ATCGGATGTCGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGCGGCGCTT 780
 DB 136604 ATCGGATGTCGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGCGGCGCTT 136545
 QY 781 GGTGATCCGAAGAAGTCTTTGGAAAGCTTTGCGCGCTCCATTTGCTTAGGATCGGTAC 840
 DB 136544 GGTGATCCGAAGAAGTCTTTGGAAAGCTTTGCGCGCTCCATTTGCTTAGGATCGGTAC 136485
 QY 841 TGTGCACTCACTGTTTCACTTCTCTCGATCACCACCTGGTGGTGTGATCTCTGG 900
 DB 136484 TGTGCACTCACTGTTTCACTTCTCTCGATCACCACCTGGTGGTGTGATCTCTGG 136425
 QY 901 TTTGTCCTTAGTTGTGTCGCCACGTTGGGTGACTTTGTTGATCGCAGTTCAAACGCGA 960
 DB 136424 TTTGTCCTTAGTTGTGTCGCCACGTTGGGTGACTTTGTTGATCGCAGTTCAAACGCGA 136365
 QY 961 TTTGGCATCAAGGATATGTCGAACCTTCTCCAGCCACCGCGGATTTGATGACCGTTT 1020
 DB 136364 TTTGGCATCAAGGATATGTCGAACCTTCTCCAGCCACCGCGGATTTGATGACCGTTT 136305
 QY 1021 GGATGTCATGCTCCCGCGCGCATGCTGACGTTGATCTCTGAGTGTGATCAGCAGTCT 1080
 DB 136304 GGATGTCATGCTCCCGCGCGCATGCTGACGTTGATCTCTGAGTGTGATCAGCAGTCT 136245
 QY 1081 GTATCGCTGTTAAGCTTTGGGCGAGCTTTAAGTTCAAAAACCTTCAAAGCGCTGAGTG 1140
 DB 136244 GTATCGCTGTTAAGCTTTGGGCGAGCTTTAAGTTCAAAAACCTTCAAAGCGCTGAGTG 136185
 QY 1141 CATAAGCTGTGACCTCAAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGGAAT 1200
 DB 136184 CATAAGCTGTGACCTCAAGCGCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACGGAAT 136125
 QY 1201 TTTTAACTGGCGAGGTTACTCAACATGCGCACGCCCAACAAAGCGGCTAATCAATG 1260
 DB 136124 TTTTAACTGGCGAGGTTACTCAACATGCGCACGCCCAACAAAGCGGCTAATCAATG 136065
 QY 1261 CACCGTATGCTCTAGTAGGAACCGATTCCGCTGG 1300
 DB 136064 CACCGTATGCTCTAGTAGGAACCGATTCCGCTGG 136025
 RESULT 3
 AX127149/c 349980 bp DNA linear PAT 11-MAY-2001
 LOCUS Sequence 7065 from Patent EP1108790.
 DEFINITION AX127149 AX114121
 ACCESSION AX127149.1 GI:14041137
 VERSION
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium Glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotides
 JOURNAL Patent: EP 1108790-A 7065 20-JUN-2001;
 KYOWA HAKKO KOGYO CO., LTD. (JP)
 FEATURES Location/Qualifiers
 source
 1..349980
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 /notes="Seq 1 to long (3,309,400) split in 11, seq 7065
 1,800,001 2,149,980"

ORIGIN

Query Match 100.0%; Score 1300; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG 60
DB 337324 GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG 337265

QY 61 GCTTAAGAAATACGTCGTTTCAAGCTGATGCGCGGGGAAAAGCAAGCTTTCTTTGCTT 120
DB 337264 GCTTAAGAAATACGTCGTTTCAAGCTGATGCGCGGGGAAAAGCAAGCTTTCTTTGCTT 337205

QY 121 GCAAGAGTGTGGAAGAAATTTTTCGAAAATGCTGCACCATCAACAGTGAATGTTA 180
DB 337204 GCAAGAGTGTGGAAGAAATTTTTCGAAAATGCTGCACCATCAACAGTGAATGTTA 337145

QY 181 GAAACTTCAAGAGAACCCATGATGAACCGGAGCAACATCACCCTCCATGAGGATGCC 240
DB 337144 GAAACTTCAAGAGAACCCATGATGAACCGGAGCAACATCACCCTCCATGAGGATGCC 337085

QY 241 CAAACCCAAAATATCGCGGTGCGAGATCTCAAGCTGCGCATTTGTGCGGATCGGACT 300
DB 337084 CAAACCCAAAATATCGCGGTGCGAGATCTCAAGCTGCGCATTTGTGCGGATCGGACT 337025

QY 301 GGGGCTCTGCTTTCTTTGGGATTTGCTTAAGCCCATGCGGTGCTGATCCTCGTTGC 360
DB 337024 GGGGCTCTGCTTTCTTTGGGATTTGCTTAAGCCCATGCGGTGCTGATCCTCGTTGC 336965

QY 361 AGTTTATGCTCAGCAACATGGAAGTGTGAGCAGACTTAAAGAGCGGCTATCA 420
DB 336964 AGTTTATGCTCAGCAACATGGAAGTGTGAGCAGACTTAAAGAGCGGCTATCA 336905

QY 421 TTTGCCACTGCGGATTTATGATCATCGCGGTGAGCAATCATCTGGTGTGATGGCCATT 480
DB 336904 TTTGCCACTGCGGATTTATGATCATCGCGGTGAGCAATCATCTGGTGTGATGGCCATT 336845

QY 481 TGGCAGATGGGCAATTTGGGCTTTTGTGGCACTGTGTTGGTGTGATGATTTCCG 540
DB 336844 TGGCAGATGGGCAATTTGGGCTTTTGTGGCACTGTGTTGGTGTGATGATTTCCG 336785

QY 541 AATTTCTACAAATCGCAAGGAAAGAGCCGCAACTATTTAGGAGCACCTCTGTGG 600
DB 336784 AATTTCTACAAATCGCAAGGAAAGAGCCGCAACTATTTAGGAGCACCTCTGTGG 336725

QY 601 CATCTTCTGCTCACCTGGATTTCAATTTTTCGGAAGCTTCGCTCGCATGCTGCTGAT 660
DB 336724 CATCTTCTGCTCACCTGGATTTCAATTTTTCGGAAGCTTCGCTCGCATGCTGCTGAT 336665

QY 661 GCAAAAACAATTCATCCCGGTGATATTTCAATTTTTCAGCTTCACTGCTGTGATCGC 720
DB 336664 GCAAAAACAATTCATCCCGGTGATATTTCAATTTTTCAGCTTCACTGCTGTGATCGC 336605

QY 721 ATCGGATGTGGCGGGTATATCGCGGTGCTTTTTCGGAAGCTTCGCAATGCGCGCGTT 780
DB 336604 ATCGGATGTGGCGGGTATATCGCGGTGCTTTTTCGGAAGCTTCGCAATGCGCGCGTT 336545

QY 781 GGTGAGTCCGAAGAGTCTTGGGAAGCTTTGCGGTTCATGTTCTTAGATCGGTAC 840
DB 336544 GGTGAGTCCGAAGAGTCTTGGGAAGCTTTGCGGTTCATGTTCTTAGATCGGTAC 336485

QY 841 TGGTGCACTGATGCTTCACTTCCTGCTGATCACCCTGCTGATGATGATGATGATGAT 900
DB 336484 TGGTGCACTGATGCTTCACTTCCTGCTGATCACCCTGCTGATGATGATGATGATGAT 336425

QY 901 TTGTGCCCTAGTTGTGCGCACGTTGGGTGACTTTGGTTGAGTCGAGTTCAACCGCA 960
DB 336424 TTGTGCCCTAGTTGTGCGCACGTTGGGTGACTTTGGTTGAGTCGAGTTCAACCGCA 336365

QY 961 TTTGGGCATCAAGGATATGTCGACCTTCTTCAGGCCACCGCGGATTTGATGACCGTTT 1020
DB 336364 TTTGGGCATCAAGGATATGTCGACCTTCTTCAGGCCACCGCGGATTTGATGACCGTTT 336305

QY 1021 GGATGCGATGCTCCCGCGCGGATGTCAGCTGGTGTGATCCTGAGTGTGATCAGCAGCTC 1080
DB 336304 GGATGCGATGCTCCCGCGCGGATGTCAGCTGGTGTGATCCTGAGTGTGATCAGCAGCTC 336245

QY 1081 GTATCGTGTGTTAAAGCTTTGGCCAGCTTTAAAGTCAAAAACTTGAAGGCGCTGAGGTG 1140
DB 336244 GTATCGTGTGTTAAAGCTTTGGCCAGCTTTAAAGTCAAAAACTTGAAGGCGCTGAGGTG 336185

QY 1141 CATACGTTGTCACCTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACGATTT 1200
DB 336184 CATACGTTGTCACCTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACGATTT 336125

QY 1201 TTCTTAACCTGGCGAGCGGTACTCAAAATGCGCACGCCACCAACAAAGCGCCATTAATCAATG 1260
DB 336124 TTCTTAACCTGGCGAGCGGTACTCAAAATGCGCACGCCACCAACAAAGCGCCATTAATCAATG 336065

QY 1261 CACCGGTAATGCTGCTAGTAGGAACCGATTCCCGCTGG 1300
DB 336064 CACCGGTAATGCTGCTAGTAGGAACCGATTCCCGCTGG 336025

RESULT 4
AX127150/c 349980 bp DNA linear PAT 11-MAY-2001
LOCUS
DEFINITION Sequence 7066 from Patent EP1108790.
ACCESSION AX127150 AX114121
VERSION AX127150.1 GI:14041138
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7066 20-JUN-2001;
KIOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
Location/Qualifiers
source
1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3,309,400) split in 11, seq 7066
2,100,001 2,449,980"

ORIGIN
Query Match 100.0%; Score 1300; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG 60
DB 337324 GAAGTCGTGCTCCCAAGGAAAAGAACTGATGAGGTCTAGAAGACCTTTATCGCAATG 337265

QY 61 GCTTAAGAAATACGTCGTTTCAAGCTGATGCGCGGGGAAAAGCAAGCTTTCTTTGCTT 120
DB 337264 GCTTAAGAAATACGTCGTTTCAAGCTGATGCGCGGGGAAAAGCAAGCTTTCTTTGCTT 337205

QY 121 GCAAGAGTGTGGAAGAAATTTTTCGAAAATGCTGCACCATCAACAGTGAATGTTA 180
DB 337204 GCAAGAGTGTGGAAGAAATTTTTCGAAAATGCTGCACCATCAACAGTGAATGTTA 337145

QY 181 GAAACTTCAAGAGAACCCATGATGAACCGGAGCAACATCACCCTCCATGAGGATGCC 240
DB 337144 GAAACTTCAAGAGAACCCATGATGAACCGGAGCAACATCACCCTCCATGAGGATGCC 337085

QY 241 CAAACCCAAAATATCGCGGTGCGAGATCTCAAGCTGCGCATTTGTGCGGATCGGACT 300
DB 337084 CAAACCCAAAATATCGCGGTGCGAGATCTCAAGCTGCGCATTTGTGCGGATCGGACT 337025

QY 301 GGGGCTCTGCTTTCTTTGGGATTTGCTTAAGCCCATGCGGTGCTGATCCTCGTTGC 360
DB 337024 GGGGCTCTGCTTTCTTTGGGATTTGCTTAAGCCCATGCGGTGCTGATCCTCGTTGC 336965

Db 37024 GGGGCTCTGGTCTTTTGGGAGTTGCTCTAAGCCATGGGTTGGTACATCTCTCGTTGC 36965
 QY 361 AGTTTATAGCTGAGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCA 420
 Db 36964 AGTTTATAGCTGAGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCA 36905
 QY 421 TTGTCACATGCGGATATGATCATCGCGGTGAGCAATCATCTGGTGTGATGCGCAAT 480
 Db 36904 TTGTCACATGCGGATATGATCATCGCGGTGAGCAATCATCTGGTGTGATGCGCAAT 36845
 QY 481 TGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTGTGATGATTTCCG 540
 Db 36844 TGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTGTGATGATTTCCG 36785
 QY 541 AATTTTCTACATGCGGAAAGAGCCGCAACTATTGAGGAGACACCTCTGTGGG 600
 Db 36784 AATTTTCTACATGCGGAAAGAGCCGCAACTATTGAGGAGACACCTCTGTGGG 36725
 QY 601 CATCTGCTGCTCACTCGATTTCCATTGTCGAGCTTCGCTCGATGCTGCTGAT 660
 Db 36724 CATCTGCTGCTCACTCGATTTCCATTGTCGAGCTTCGCTCGATGCTGCTGAT 36665
 QY 661 GCAAAACAAATTCATCCCGGTGATATTCATTTGAGCTTCATGCTGTGATGCTG 720
 Db 36664 GCAAAACAAATTCATCCCGGTGATATTCATTTGAGCTTCATGCTGTGATGCTG 36605
 QY 721 ATCGGATGCGCGGTATATCGCGGTGCTCTTTGGATCGCACCAATGCGCGCTT 780
 Db 36604 ATCGGATGCGCGGTATATCGCGGTGCTCTTTGGATCGCACCAATGCGCGCTT 36545
 QY 781 GGTGAGTCCGAAGAGTCTTGGGAAGCTTTGCGGCTCCATTGCTTAGATCGGTAC 840
 Db 36544 GGTGAGTCCGAAGAGTCTTGGGAAGCTTTGCGGCTCCATTGCTTAGATCGGTAC 36485
 QY 841 TGGTCACTCACTGTTCACTTCTCTGCTGATCACCACCTGTTGGTGTGATCTTGGG 900
 Db 36484 TGGTCACTCACTGTTCACTTCTCTGCTGATCACCACCTGTTGGTGTGATCTTGGG 36425
 QY 901 TTGTCCTCTAGTTGTTGCGGCACTGTTGGGTGACTTGGTGGTGGTGGTGGTGGT 960
 Db 36424 TTGTCCTCTAGTTGTTGCGGCACTGTTGGGTGACTTGGTGGTGGTGGTGGTGGT 36365
 QY 961 TTGTCCTCTAGTTGTTGCGGCACTGTTGGGTGACTTGGTGGTGGTGGTGGTGGT 1020
 Db 36364 TTGTCCTCTAGTTGTTGCGGCACTGTTGGGTGACTTGGTGGTGGTGGTGGTGGT 36305
 QY 1021 GGATGGCATGCTCCCGGCGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080
 Db 36304 GGATGGCATGCTCCCGGCGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 36245
 QY 1081 GTATCCGTCGTAAGCTTGGGCGGATTTAGTTTCAAAACCTTTGAAGGCGCTGAGTG 1140
 Db 36244 GTATCCGTCGTAAGCTTGGGCGGATTTAGTTTCAAAACCTTTGAAGGCGCTGAGTG 36185
 QY 1141 CATAACGTGTGACCTCAGCGCTTTTGGTGTCTCAAAAGTTTAAAGGCGCTTACGGATT 1200
 Db 36184 CATAACGTGTGACCTCAGCGCTTTTGGTGTCTCAAAAGTTTAAAGGCGCTTACGGATT 36125
 QY 1201 TTCTTAATGCGGAGGCTTCAACATGCGCAGCCACCAAGGCGCATATCAATG 1260
 Db 36124 TTCTTAATGCGGAGGCTTCAACATGCGCAGCCACCAAGGCGCATATCAATG 36065
 QY 1261 CACCGGTAATGGCTAGTAGGAAACCGATTTCGGGTGG 1300
 Db 36064 CACCGGTAATGGCTAGTAGGAAACCGATTTCGGGTGG 36025

RESULT 5
 LOCUS CGL312009/c 1194 bp DNA linear BCT 30-JAN-2002
 DEFINITION Corynebacterium glutamicum partial cdsA gene for CDP-diacylglycerol
 synthase.
 ACCESSION AJ312009

VERSION AJ312009.1 GI:16505727
 KEYWORDS CDP-diacylglycerol synthase; cdsA gene.
 SOURCE Corynebacterium glutamicum
 ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 Nampoothiri, K.M., Hoischen, C., Bathe, B., Mockel, B., Pfeifferle, W., Krumbach, K., Sahm, H. and Eggeling, L.
 AUTHORS Expression of genes of lipid synthesis and altered lipid composition modulates L-glutamate efflux of Corynebacterium glutamicum
 TITLE Appl. Microbiol. Biotechnol. 58 (1), 89-96 (2002)
 JOURNAL 21820045
 MEDLINE 11831479
 PUBMED 2 (bases 1 to 1194)
 REFERENCE Eggeling, L.
 AUTHORS Direct Submission
 TITLE Submitted (25-APR-2001) Eggeling L., Biotechnology,
 JOURNAL Forschungszentrum, Forschungszentrum, D-52425 Juelich, GERMANY
 FEATURES
 Location/Qualifiers
 1..1194
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 1..1194
 /gene="cdsA"
 <1..>1194
 /gene="cdsA"
 /product="CDP-diacylglycerol synthase"

ORIGIN
 Query Match 91.8%; Score 1193; DB 1; Length 1194;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAGTCGTCTCGCAGAAAGCACTGATGGAGTCTAGACACCTTATCGCAATG 60
 Db 1193 GAAGTCGTCTCGCAGAAAGCACTGATGGAGTCTAGACACCTTATCGCAATG 1134
 QY 61 GCTTAAGAAATACGTCGTTTCAACGTCGATTTGGCGGGGAAACGACGCTTTCTTTGCTT 120
 Db 1133 GCTTAAGAAATACGTCGTTTCAACGTCGATTTGGCGGGGAAACGACGCTTTCTTTGCTT 1074
 QY 121 GCAAGAGTGTTCGAAAGATTTTTCGAAATGCTGGCACCACATCAACAGTGACATTGTTA 180
 Db 1073 GCAAGAGTGTTCGAAAGATTTTTCGAAATGCTGGCACCACATCAACAGTGACATTGTTA 1014
 QY 181 GAAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 240
 Db 1013 GAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCC 954
 QY 241 CAAACCCAAAATAATGCGGTCGAGATCTCAAGCTGCCATTTGCTGTGGGATCGGACT 300
 Db 953 CAAACCCAAAATAATGCGGTCGAGATCTCAAGCTGCCATTTGCTGTGGGATCGGACT 894
 QY 301 GCGGGTCCTGGTCTTTTGGGATTTGCTAAGCCCATGGGTTGTTACATCTCGTTGC 360
 Db 893 GCGGGTCCTGGTCTTTTGGGATTTGCTAAGCCCATGGGTTGTTACATCTCGTTGC 834
 QY 361 AGGTTTATGGCTGAGCAACATGGAAGTTGGTAGAGACTTAAAGAGCGGCTATCA 420
 Db 833 AGGTTTATGGCTGAGCAACATGGAAGTTGGTAGAGACTTAAAGAGCGGCTATCA 774
 QY 421 TTTGCCACTGCGGATATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCCATT 480
 Db 773 TTTGCCACTGCGGATATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCCATT 714
 QY 481 TGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGGCACTGTGTTGGCTGATTTCCG 540
 Db 713 TGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGGCACTGTGTTGGCTGATTTCCG 654
 QY 541 AATTTTCTACATGCGCAGGAAAGAGCCGCAACTATTGAGGAGACACCTCTGTGGG 600

653 AATTTTCTACAAATGGCCGGAAGAAAGCCGCAACTATTTGAGGACACCTCTGTGGG 594
601 CATCTTCTGCTCACCTGGGATTCATTTGTCGGAAGCTTCGCTCGGATGCTGTGCTGAT 660
593 CATCTTCTGCTCACCTGGGATTCATTTGTCGGAAGCTTCGCTCGGATGCTGTGCTGAT 534
661 GCAAAACAAATTCATCCCGGGTACATATTTTCATTTGACGTTTCATGCTGTGTGATCGC 720
533 GCAAAACAAATTCATCCCGGGTACATATTTTCATTTGACGTTTCATGCTGTGTGATCGC 474
721 ATCCGATGTCGGGGGATATCCGGGGTGTCTTCTTTGGATCCGACCCCAATGGCGCGGTT 780
473 ATCCGATGTCGGGGGATATCCGGGGTGTCTTCTTTGGATCCGACCCCAATGGCGCGGTT 414
781 GGTGAGTCGGAAGATCTTTGGGAAGCTTTTGGCGGCTCCATGTTCTTAGGATCGGTAC 840
413 GGTGAGTCGGAAGATCTTTGGGAAGCTTTTGGCGGCTCCATGTTCTTAGGATCGGTAC 354
841 TGGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCCTGCTGGATGGTGTGATCTTTGG 900
353 TGGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCCTGCTGGATGGTGTGATCTTTGG 294
901 TTGTGCGCTAGTTGTGTCGCGACGTTTGGGTGACTTGGTTGAGTCGCGAGTTCAAACGCGA 960
293 TTGTGCGCTAGTTGTGTCGCGACGTTTGGGTGACTTGGTTGAGTCGCGAGTTCAAACGCGA 234
961 TTGTGGGATCAAGGATATGTCGACCTTCTTCCAGGCCACCGGGATTTGATGACGCTC 1020
233 TTGTGGGATCAAGGATATGTCGACCTTCTTCCAGGCCACCGGGATTTGATGACGCTC 174
1021 GGATGGGATGCTCCCGGCGGATGTTGACGTTGCTGATCTGAGTGTGATCAGCAGCTC 1080
173 GGATGGGATGCTCCCGGCGGATGTTGACGTTGCTGATCTGAGTGTGATCAGCAGCTC 114
1081 GTATCCGTCGTAAGCTTTGGGCGGATTTAAGTTCAAAAACCTTGAAGGCGCTGAGGTG 1140
113 GTATCCGTCGTAAGCTTTGGGCGGATTTAAGTTCAAAAACCTTGAAGGCGCTGAGGTG 54
1141 CATACGCTGTCACCTCAGCGGCTTTTGGGCTGTCGAAAGTTTAAAGGCTTTA 1193
53 CATACGCTGTCACCTCAGCGGCTTTTGGGCTGTCGAAAGTTTAAAGGCTTTA 1

RESULT 6
AX064835
LOCUS AX064835 978 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 1117 from Patent WO0100843.
ACCESSION AX064835
VERSION AX064835.1 GI:12542547
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1
Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.
Corynebacterium glutamicum genes encoding metabolic pathway
proteins
JOURNAL patent: WO 0100843-A 1117 04-JAN-2001;
BASF AKTIENGESSELLSCHAFT (DE)
FEATURES
Location/Qualifiers
1..978
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101..958
/note="unnamed protein product; RXA01894"
/codon_start=1
/transl_table=11
/protein_id="CAC25657.1"
/db_xref="GI:12542548"
/db_xref="REMYEMBL:CAC25657"

/translation="MPKPNAGRLDKAAIAVIGIGLVLLGIVLSPGWVILVAGF
MAATWVGSRLEKGGVHLPIPLIMLIGQALIMLSPFGTMIILASFAVATVLMVFR
IFYNGTEKARNYLRDTSVGLFVLTWPLPSPFAAMLSLMQNNISPGYFTLTLVLCV
IASDVGVIAGVFFGSGMAPLSPFKSGVAGSIVLGSVGTALSHFLDHHWMS
VILGCALVVCATLGLDLESQFKRLDGIKDNLSLFLPGHGLMDRLDGLPAAAVTLIL
SVISSSSYS"

Query Match 75.2%; Score 978; DB 6; Length 978;
Best Local Similarity 100.0%; Pred. No. 1.2e-295; Indels 0; Gaps 0;
Matches 978; Conservative 0; Mismatches 0;

ORIGIN

QY 136 AGAATTTTTTTCGAAATAGTGTGCACCATCAACAGTACATTTTAAACACTTCAAGGAGA 195
DB 1 AGAATTTTTTTCGAAATAGTGTGCACCATCAACAGTACATTTTAAACACTTCAAGGAGA 60
QY 196 ACCATGAATGAACCGGAGCAACATCACGGTCCATGAGGATGCCCAACCCAAATAA 255
DB 61 ACCCATGAATGAACCGGAGCAACATCACGGTCCATGAGGATGCCCAACCCAAATAA 120
QY 256 TCGGGTTCGAGATCTCAAGAGTCCCATTCGCTGTGGGATCGGACTGGGGTCTCTGTTCT 315
DB 121 TCGGGTTCGAGATCTCAAGAGTCCCATTCGCTGTGGGATCGGACTGGGGTCTCTGTTCT 180
QY 316 TTTGGGATTTGCTTAAGCCCATGGGGTTGGTACATCTCTGTTGCAAGTTTATGCTCC 375
DB 181 TTTGGGATTTGCTTAAGCCCATGGGGTTGGTACATCTCTGTTGCAAGTTTATGCTCC 240
QY 376 ACNACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 435
DB 241 AGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 300
QY 436 TATGATCATCGGGGTTCAGGCAATCATCTGGTGTCTATGCGCATTTGGCACGATGGCAT 495
DB 301 TATGATCATCGGGGTTCAGGCAATCATCTGGTGTCTATGCGCATTTGGCACGATGGCAT 360
QY 496 TTTGGGCTCTTTTGGCCACCTGTGTTGGTGTGATTTTCGAAATTTCTCAATCG 555
DB 361 TTTGGGCTCTTTTGGCCACCTGTGTTGGTGTGATTTTCGAAATTTCTCAATCG 420
QY 556 CACGAAAAAGAGCCCGCAACTATTTAGAGGACACCTCTGTGGGATCTCTGTGCTAC 615
DB 421 CACGAAAAAGAGCCCGCAACTATTTAGAGGACACCTCTGTGGGATCTCTGTGCTAC 480
QY 616 CTGATTCATTTTCGGAAGCTTCGCTCGGATGCTGTGCTGATGCAAAACAATTCAT 675
DB 481 CTGATTCATTTTCGGAAGCTTCGCTCGGATGCTGTGCTGATGCAAAACAATTCAT 540
QY 676 CCGGGTACATATTTTCATTTTTCAGTTTCATGTTGATCGCATCGGATGTTGGCGG 735
DB 541 CCGGGTACATATTTTCATTTTTCAGTTTCATGTTGATCGCATCGGATGTTGGCGG 600
QY 736 GTATATCGGGGTGTGTTCTTTTGATTCGACCAATGGCGCGGTTGGTGGTCCGAAGAA 795
DB 601 GTATATCGGGGTGTGTTCTTTTGATTCGACCAATGGCGCGGTTGGTGGTCCGAAGAA 660
QY 796 GTCTTGGGAAGCTTTTGGCGGCTCCATTCCTTTAGGATCGGCTGCTGCTCAGTGT 855
DB 661 GTCTTGGGAAGCTTTTGGCGGCTCCATTCCTTTAGGATCGGCTGCTGCTCAGTGT 720
QY 856 TCACCTTCCTGCTCGATCACCCTGTTGGTGTGATCTTTGGGTTGTGCCCTAGTTGT 915
DB 721 TCACCTTCCTGCTCGATCACCCTGTTGGTGTGATCTTTGGGTTGTGCCCTAGTTGT 780
QY 916 GTGCGCACCTTCTTCCAGGCCACCGGATTTGATGAGCCGTTTGGATGGCATGTCTCC 1035
DB 781 GTGCGCACCTTCTTCCAGGCCACCGGATTTGATGAGCCGTTTGGATGGCATGTCTCC 840
QY 976 TATGTCGAACCTTCTTCCAGGCCACCGGATTTGATGAGCCGTTTGGATGGCATGTCTCC 1035
DB 841 TATGTCGAACCTTCTTCCAGGCCACCGGATTTGATGAGCCGTTTGGATGGCATGTCTCC 900
QY 1036 GGCCGCGATGCTGACGTGGTGTGATCTCTGATGTGATCAGCAGCTCGTATCCGTCGTAAG 1095


```
QY 236 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGTGTGGGATC 295
Db 1 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGTGTGGGATC 60
QY 296 GGAATGGGGTCCTGGTTCCTTTGGGGATTTGCTTAAGCCCATGGGTTGTATCATCTC 355
Db 61 GGAATGGGGTCCTGGTTCCTTTGGGGATTTGCTTAAGCCCATGGGTTGTATCATCTC 120
QY 356 GTTCAGGTTTATGGCTGCGAGCAATGGAAGTTGGTAGCAGACTAAAGAGGGCG 415
Db 121 GTTCAGGTTTATGGCTGCGAGCAATGGAAGTTGGTAGCAGACTAAAGAGGGCG 180
QY 416 TATCATTTGGCACTGCCGATATGATCATCGCGGGTCAGGCAATCATCTGCTCATGG 475
Db 181 TATCATTTGGCACTGCCGATATGATCATCGCGGGTCAGGCAATCATCTGCTCATGG 240
QY 476 CCATTTGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTCTGATGAT 535
Db 241 CCATTTGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTCTGATGAT 300
QY 536 TTCCGAATTTTCTACAATGCGACGAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 595
Db 301 TTCCGAATTTTCTACAATGCGACGAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 360
QY 596 GTGGGCAATCTTCTGCTCACTCGGATTCATGTTGGAGCTTCGCTGCGATGCTGCG 655
Db 361 GTGGGCAATCTTCTGCTCACTCGGATTCATGTTGGAGCTTCGCTGCGATGCTGCG 420
QY 715 CTGATGCAAAACAATTCATCCCGGGTACATATTTGAGTTTCATGCTGTGTTG 715
Db 421 CTGATGCAAAACAATTCATCCCGGGTACATATTTGAGTTTCATGCTGTGTTG 480
QY 716 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTGTTCTTTGGATCGCACCACCAATGGCG 775
Db 481 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTGTTCTTTGGATCGCACCACCAATGGCG 540
QY 776 CCGTTGGTGAATCGAAGAACTCTTGGAGGCTTTGCGCGCTCCATGCTTAGGATCG 835
Db 541 CCGTTGGTGAATCGAAGAACTCTTGGAGGCTTTGCGCGCTCCATGCTTAGGATCG 600
QY 836 GTCACTGTGTGACACTCAGTTGTTCACTTCTGCTCGATCACCACTGGTGGATGGTGTGATC 895
Db 601 GTCACTGTGTGACACTCAGTTGTTCACTTCTGCTCGATCACCACTGGTGGATGGTGTGATC 660
QY 896 TTGGGTTGTGCCCTAGTTGTGTGCCACGCTTGGGTTGAGTTGAGTCCAGTTCAAA 955
Db 661 TTGGGTTGTGCCCTAGTTGTGTGCCACGCTTGGGTTGAGTTGAGTCCAGTTCAAA 720
QY 956 CGGATTTGGGCATCAGGATATGCGAACCTTCTCCAGGCCACGGGATTCATGGAC 1015
Db 721 CGGATTTGGGCATCAGGATATGCGAACCTTCTCCAGGCCACGGGATTCATGGAC 780
QY 1016 CGTTTGGATGATGCTCCCGCGCGGATGTTGACGTTGTTGATCCTGAGTTGATCAGC 1075
Db 781 CGTTTGGATGATGCTCCCGCGCGGATGTTGACGTTGTTGATCCTGAGTTGATCAGC 840
QY 1076 AGTCGGATCCGTCG 1090
Db 841 AGTCGGATCCGTCG 855
RESULT 9
BD16417
LOCUS BD16417 855 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD16417
VERSION BD16417.1 GI:27870229
KEYWORDS JP 2002191370-A/2216.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 855)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
```

```
Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
Novel polynucleotide
Patent: JP 2002191370-A 2216 09-JUL-2002;
KYOWA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/2216
PD 09-JUL-2002
PP 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
PI OZAKI
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
PC 04, C12P13/08,
PC C12P15/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/566,
PC G01N33/569, G01N33/68, G01N37/00, C12P21/08, C12N1/21, C12R1/15,
PC (C12N1/21, C12R1/13), (C12N1/21, C12R1/01), (C12P13/08, C12R1/15),
PC C12N15/00,
PC C12N5/00, C12N15/00
CC Novel polynucleotide
PH Key Location/Qualifiers
FT source i. .855 /organism='Corynebacterium glutamicum'.
FEATURES
source
1. .855
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 65.8%; Score 855; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 5.8e-257;
Matches 855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGTGTGGGATC 295
Db 1 ATGCCAAACCCAAAAATAATGCGGGTCGAGATCTCAAAAGCTGCCATTGTGTGGGATC 60
QY 296 GGAATGGGGTCCTGGTTCCTTTGGGGATTTGCTTAAGCCCATGGGTTGTATCATCTC 355
Db 61 GGAATGGGGTCCTGGTTCCTTTGGGGATTTGCTTAAGCCCATGGGTTGTATCATCTC 120
QY 356 GTTCAGGTTTATGGCTGCGACCAATCGGAAGTTGGTAGCAGACTTAAAGAGGGCGC 415
Db 121 GTTCAGGTTTATGGCTGCGACCAATCGGAAGTTGGTAGCAGACTTAAAGAGGGCGC 180
QY 416 TATCATTTGGCACTGCCGATATGATCATCGCGGGTCAGGCAATCATCTGCTCATGG 475
Db 181 TATCATTTGGCACTGCCGATATGATCATCGCGGGTCAGGCAATCATCTGCTCATGG 240
QY 476 CCATTTGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTCTGATGAT 535
Db 241 CCATTTGGCAGATGGGCAATTTGGGCTCTTTTGGGCACTGTGTTGGTCTGATGAT 300
QY 536 TTCCGAATTTTCTACAATGCGACGAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 595
Db 301 TTCCGAATTTTCTACAATGCGACGAAAAAGAACCCCGCAACTATTTGAGGGACACCTCT 360
QY 596 GTGGGCAATCTTCTGCTCACTCGGATTCATGTTGGAGCTTCGCTGCGATGCTGCG 655
Db 361 GTGGGCAATCTTCTGCTCACTCGGATTCATGTTGGAGCTTCGCTGCGATGCTGCG 420
QY 715 CTGATGCAAAACAATTCATCCCGGGTACATATTTGAGTTTCATGCTGTGTTG 715
Db 421 CTGATGCAAAACAATTCATCCCGGGTACATATTTGAGTTTCATGCTGTGTTG 480
QY 716 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTGTTCTTTGGATCGCACCACCAATGGCG 775
Db 481 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTATATCGCGGGTGTGTTCTTTGGATCGCACCACCAATGGCG 540
```



```
QY 776 CCGTTGGTGAGTCCGAAGAAGCTTTGGGAAGGCTTTGCCGGTCCATCTTCTAGGATCG 835
Db 541 CCGTTGGTGAGTCCGAAGAAGCTTTGGGAAGGCTTTGCCGGTCCATCTTCTAGGATCG 600
QY 836 GTCACTGGTGCACTCAGTGTCTTCACTTCTCTGCTCGATACCACTCGTGATCGGCTGTGATC 895
Db 601 GTCACTGGTGCACTCAGTGTCTTCACTTCTCTGCTCGATACCACTCGTGATCGGCTGTGATC 660
QY 896 TTGGCTTGGCCCTAGTGTGTGGCCACAGTGTGGGTGCTTGTGAGTCCAGTTCACAA 955
Db 661 TTGGGTGTGGCCCTAGTGTGTGGCCACAGTGTGGGTGCTTGTGAGTCCAGTTCACAA 720
QY 956 CCGCATTTGGGCATCAAGGATATGTCGAACCTCTTCCAGGCCACGGCGGATTCATGAC 1015
Db 721 CCGCATTTGGGCATCAAGGATATGTCGAACCTCTTCCAGGCCACGGCGGATTCATGAC 780
QY 1016 CTTTGGATGGATGCTCCCGCCCGGATGTGACGTGTGATCCTGATGATGATCAGC 1075
Db 781 CTTTGGATGGATGCTCCCGCCCGGATGTGACGTGTGATCCTGATGATGATCAGC 840
QY 1076 AGCTCGTATCCGTCG 1090
Db 841 AGCTCGTATCCGTCG 855

RESULT 10
AP005220/c
LOCUS
DEFINITION
  Corynebacterium efficiens YS-314 DNA, complete genome, section
  7/11.
AP005220 BA000035
VERSION
  AP005220.1 GI:23493593
KEYWORDS
  Corynebacterium efficiens YS-314
  Corynebacterium efficiens YS-314
  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
  Corynebacterineae; Corynebacteriaceae; Corynebacterium.
ORGANISM
  1
  Nishio,Y., Nakamura,Y., Kawarabayasi,Y., Usuda,Y., Kimura,E.,
  Sugimoto,S., Matsui,K., Yamagishi,A., Kikuchi,H., Ikeo,K. and
  Gojobori,T.
  Comparative Complete Genome Sequence Analysis of the Amino Acid
  Replacements Responsible for the Thermostability of Corynebacterium
  efficiens
  Genome Res. 13 (7), 1572-1579 (2003)
  2 (bases 1 to 303250)
  12840036
  Kawarabayasi,Y., Yamazaki,J., Hino,Y., Kikuchi,H. and
  Director-General of Biotechnology Center.
  Direct Submission
  Submitted (17-MAY-2002) Director-General of Biotechnology Center,
  National Institute of Technology and Evaluation, Biotechnology
  Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan
  (E-mail: bikenite.go.jp, Tel:81-3-3481-1933, Fax:81-3-3481-8424)
  Kawarabayasi, Y. is officially affiliated with the National
  Institute of Advanced Industrial Science and Technology, Tsukuba,
  Ibaraki, 305-8566 Japan
  Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the
  National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
  Itoh, T. is at the Japan Biological Information Research Center,
  Koto-ku, Tokyo, 135-0064 Japan
  Yamagishi, A. is at Tokyo University of Pharmacy and Life Science,
  Hachioji, Tokyo, 192-0392 Japan
  Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co.,
  Inc., Kawasaki, Kanagawa, 210-8681 Japan
  The other authors are at the National Institute of Technology and
  Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.
  Location/Qualifiers
  1..303250
  /organism="Corynebacterium efficiens YS-314"
  /mol_type="genomic DNA"
  /strains="YS-314"
```

```
/db_xref="taxon:196164"
60..941
/note="CB1753, similar to AB007099-6|AAK46964.1| percent
identity: 46 in 283 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18563.1"
/db_xref="GI:23493594"
/transl_table=11
/note="WTFRGDISRSGKARGGGRGMVGGGLGGLLVLGLVLLGN
PAQVELLQGGQQLQLETSFCHQCTGEDANTYDDCLRYTMSVNEWQVLPQ
ABQYSEPELELFTNTTSGCCASAAATPFYCPADQAYFDLTTFDQMRQFGGAP
FAQMYTVAEVEHGVHONLEGLTGLYNDPGASNAVRLQADCYAGIWAYADKGE
DALLFPVTEELDAALTAAAVGDDNIQRRSGGEVNPESWTHSGSEQRQAFLDGYTT
GOMSSCDYLDRGVYRTT"
complement (938..3598)
/note="CB1754, similar to AB056583-2|BAB64333.1| percent
identity: 34 in 708 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18564.1"
/db_xref="GI:23493595"
/transl_table=11
/note="NDRTHLOTQTAQOHLNRAQAOEVLRLSPARVHHRKIGF
GPDQSRVILRDADLDSANFYTRVALDGAERQLRYLTATAVLPGLMETARAGT
GGFLRRFPESIRRGEALADLTTRVADLDKQFGLTPLDITSTATTQGGVHLF
PGAHVQRYIDARAEIARQLEDGAPGFWQDGHHLASVHRLVHTYASDNPQVRLR
AAERALSIGNQRVEILLEQLPVDALEATNRLRFLGLDITGISTVADVLRASESL
LATAQGHOTARHKMAAQTLQEQALGRNTVLIGDTPQAVALIRILAEFEQADAL
SPAERHRRVDIYIQLPATISATAGTVVVGAGPDGQDAWNRIITDGLAWVSNPN
LHFQHIPTPPADIHDDVITRPAHYQGLLTLGREIEGVDEHLDITVLQIRELEL
TTHRLDLHURGQSCGAFPTVVKVLLGDDMGLKTIQAAVAHRLTAEDTFTLV
VPSAVIWNVEECRRFTDLPVFAHQDQKAAVTAWADTGILICTTAGTRSLDIPA
VGLIITADEAHMKNPQSRQCARLIDAAVALMTGTPLENKYSFVALIRYVQPE
LITRGSAMSAGDFRARIAPAYLRKNQTDVLDDELDESDIMELTADPADVDDQV
RAGNMWGRSADGLAPKPSATSKMKRIIELLDEAEHGRKTLITFLDLILDEHV
LCAVRTIGDQSVPAARQELVDALADAPGASALISQITAGVGNINQASQVLTCEP
QVKPTIEQAVARVHRMQGTATVNHRLIGDITADERMLEILAGTKQIFDYVARLSDI
AGTPDAVDITEAQLAASIIDAEARLGLD"
3775..4158
/note="CB1755"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC18565.1"
/db_xref="GI:23493596"
/transl_table=11
/note="MPALDPHAPPELKHQKOREATMKDTHGCHDDDAPPPRRRESYR
SRDIQVITDEDDGISLELDGTPVEVYVNGKYYSPTAHMYVEFDLTDEVVDLIRTS
DEYWLHPPEDDGPNPHRSPTGE"
4162..5409
/note="CB1756, similar to M11302-1|AAA26834.1| percent
identity: 45 in 295 aa"
/codon_start=1
/transl_table=11
/product="putative tyrosinase"
/protein_id="BAC18566.1"
/db_xref="GI:23493597"
/transl_table=11
/note="WHIRQNIYALDEQLERFQDAVNGIKADGYDHYFTEOHHHSMHE
ATVFWESGHLRLNSAHRGAPLFWHRYCYREPELALQEKDPSVPLPYWDAWDSAA
GAAPLWNTDPRRIYIGNGGTGDSVNTGPFAGVSVLVWGMGFRFRGGLRQL
GLAVPTLPFDDEVATCIDLDPVYDTPWHPGSADSFNQLGPNWGPAMHNRVHWVG
GMVGTGPNDFVFLYHHAFVLDIWARQQTHAGGLPDGGGPGPHNLHDEMYLDST
SPTPANCIDYRHSMGYIYDTPPTVQCTTALDFTNVMAGDVERRYALFHARAQPV
FVPLTRGLQAPYSLFTPEPTIIEHPVDNRPYDEVRIWFATLDTLPGPAPAGSAQV
RCRETDEIFSLTGSTRELSVR"
5411..5229
/note="CB1757, similar to AB007745-6|AAK80407.1| percent
identity: 22 in 221 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18567.1"
/db_xref="GI:23493598"
```



```
/translation="MTPTPTPRGASCWGARGVEKGPWLQWDMGARDTTPADPAVA
PLASNDIEAHLRSQLNKAKSDSFRWVNRVLAERKHLISAQSSALDAFAELRNASHG
PYNLRIPADRPDTVAMI OHIRDLDDPFALEVLISHQVRSIQVLDQDPPIAAALKIIR
TEISQPVVDGSGYVALLTNTIARWAADSDNSILDASTVAEVLDAETSDSAVF
MPRTATQAEVDVNTGPELPSPVIVTEHGKPSQKPLRLISGRDMRSLIELVELR"
CDS
6300..7421
/note="CE1758, similar to AL049707-13|CAB41282.1| percent
identity: 73 in 356 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18568.1"
/db_xref="GI:23493599"
/translation="MQRGIFITGIDVTPTNGTPTGEGIRATWOLIAKAEVGLDV
PATGEHNPFVSSPTTHLAYIAAOPTNLLI STATLITNDPVIKAEYALQHLIS
GVRDLMMGRGNTGPTVIFWPKGDIRQIPLAENIHLLRWEDVVMGKRYLPLQ
GYSTPAPLDGVAFPVHWSIRSTEIAOAFYDGFHNNIFWNKEHIAQWNLFLYQ
RFEYIGHRADQATVGGGVFGIDTESQAKDFRPYFNAPYHGSPMEDTRMTP
LTGTAEQVTERTEFPADWVDYQORQLFDHAGLPLEVVLDOIERLGRDVPVRRR
MEERRPDHVPSPDPTHTLLGNPEHPHFHISPGPADAE"
7444..8121
/note="CE1759, similar to AL049707-12|CAB41281.1| percent
identity: 52 in 185 aa"
/codon_start=1
/transl_table=11
/product="putative oxidoreductase"
/protein_id="BAC18569.1"
/db_xref="GI:23493600"
/translation="MRLLTVVTAGLSNPFSTRSVADOLTNAVEAATARGESVDIEVI
EIRLIIFDLATFTAGMSFPALDPAKERAHSDGLIATVPVTASVSGIFKMPFPAI
DKPTVIGLPLTIIAASAGTERHSLVDHAIKPMNYLEAVVPTGVAATAATDFAGV
EFERRVNRAGELATLMQDYSVOGLGGATQADADLSYRTGTVPGEFTSFADIL
KGHDGGO"
CDS
complement(8359..9042)
/note="CE1760, similar to AE000498-8|AAC77234.1| percent
identity: 27 in 203 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC18570.1"
/db_xref="GI:23493601"
/translation="MRSRLCLFVFKGRYTPADLTTLTDELIDRLDPGSLVADRGF
YSFRWATQAGRDALVRIKNNVKKPHVRLTEGTLWALITPSNGPQVAPRMVR
VVDYHIDTGDDEAGIVQOHLRLTLDPDEVSVEDLIAAYHTREWLENTFDEBKTHQ
RGRERVLRSKSPVLRQIEVHLCCHYAIRTLMAATAASEHPVRSFVYKALKLTR
SVIQGSPPP"
CDS
complement(9074..10327)
/note="CE1761, similar to AF024666-31|AA033382.1| percent
identity: 84 in 408 aa"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAC18571.1"
/db_xref="GI:23493602"
/translation="MAAGPVSIDETTYLDELLAASPLMREVMLOGFINOILSTQADQ
VCGADIARVQDRVNRVYRHRDFTRVGTIDVAVPKLRTGTFPPWILERRTRABR
ALTVIATCYLKGYSRTRMDLVATLGINNLSQSDMDALDDDDWEEBFRPLDTG
PLYVSCDALTMKVRGEGRVKTSVILLATGVNAEYRLLGMQVATAESVSWTGFPR
DLKARGNEVLVTSDAHGLGQHAIGVLPNASWQRCTHFMSKNLYGMSKQWTPILS
AMPTFIPOQDAQVWGQARVAVFCEQKPDVADYLEALDELIIAFTNAPKSVWTKV
WNNPTELRNEIRRRRTDVVGIFPNRDVAVRLVGAVLAEQHDDWIQQKRYMSLTSLEQ
TRAMTANILSDDIITSEIRKDYA"
complement(10434..11051)
/note="CE1762"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC18572.1"
/db_xref="GI:23493603"
/translation="MPRHGRTPQDTRRLSDLSISVGLLTVRFVADVVDVDEVAQGRQV
QRSRLSPRVMYVITWALNSDGSYEDIADLTIDGLSWACRWEQVAPFAPSATFQ
ARVLGHEPMAALFARVAQPARFEGEWAGLAKRLLAVDGTITLDDVPTTANVHFGR
PQTGRGQAGYPOARVAIAECGSHAILDAVIGGLGTGETTIGGVH"
```

```
CDS
complement(11090..12379)
/note="CE1763, similar to AL357523-32|CAB93388.1| percent
identity: 55 in 418 aa"
/codon_start=1
/transl_table=11
/product="putative histidyl-tRNA synthetase"
/protein_id="BAC18573.1"
/db_xref="GI:23493604"
Query Match 45.9%; Score 596.8; DB 1; Length 303250;
Best Local Similarity 67.7%; Pred. No. 4.2e-175;
Matches 884; Conservative 0; Mismatches 412; Indels 10; Gaps 3;
QY 1 GAAGTCGTTGCTCGCAAGAAAAGAACTGATGGAGGTCTAGAACACCTTTATCGCAATG 60
DB 173786 GAGATTGTTGCCCGAGAGAAAGCCGAGTTGATGGAGGCTAGAA--GCCCCCATCATCCAG 173729
QY 61 GCTTAAGAAATTACGTGTTTCAACGTGATGGCGGGGAAAACGACGCTTCTTTTTCGTT 120
DB 173728 GAGTCGTCGTTTCGTGCG--TCACAGTCAGTGGCGGGAGCGGCGTTCCTTCTGTTCC 173671
QY 121 GCAAGAGTCTTTGGAGAGATTTCGAAAAATGCTGGCAACCATCAACAGTGACATTGTTA 180
DB 173670 GTGTGCCCGGTGATACCCCTGGGATGCCGATGACCGGATCTGCGTTTCATCAGTTT 173611
QY 181 GAAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGCC 240
DB 173610 AAACCCGAGGAGAACCCCATGATGAGCGCGAGACCATCACCGACCGTGGGATGCC 173551
QY 241 CAACCCAAAATATCGGGTCCGATCTCAAGCTGCATTCGTCGGGATCGGACT 300
DB 173550 CAACCCAAAATATGTCGGGTGCGTATCTCAAGCGGGCGGCTCGGCGTGGGTGTCGGTCT 173491
QY 301 GGGGGTCTCTGGTTCCTTTTGGGGATTGCTCAAGCCCATGCGGGTGGTATACATCCTCGTTGC 360
DB 173490 CGGACTCTCTGTCATCGTCGGCATCCTACTACGCCCTCGGGCTGGTACCTCGTTGTCGC 173431
QY 361 AGTTTTTATGGTGCAGCAACATGGGAAGTTGGTAGCAGACTTAAGAAGCGGCTATCA 420
DB 173430 CGGCGCTGTTTGGCGTGGCGACCTGGGAGGTTTGTGGCGGCTCAAGGAAGCGGTTACCA 173371
QY 421 TTTGCCACTGCCGATTATCATCATCGCGCGTTCAGGCAATCATCTGGCTGTCAATGGCAAT 480
DB 173370 CTTGCCCATGCCATCATGATCATCGTGGGAGGCCATGGTGTGTGTTCTTCTGCGCGTT 173311
QY 481 TGGCAGCATGGGCATTTTGGCGTCTTTTGTGCCACTGTGTTGGTGTCTGATGATTTTCGG 540
DB 173310 TGAGACCTCCGGGATCCTTGGCCTCCTATGTGGCACCGCTTGTGGTGTGATGTTCTTCGG 173251
QY 541 AATTTTCTACATGGCACCGAAAAGAACCCCGCAACTATTTGAGGGACACCTCTGTGGG 600
DB 173250 GATCTTTCTACAACGGCGCCCGAGAGAGGAGCCAGGAACCTACCTGCGGGATACCTCAGTGG 173191
QY 601 CATCTTTCGTGCTCACCTGGATTCCATTGTTGCGAAGCTTCGTTGCGATGCTGTGCGTGTAT 660
DB 173190 TATCTTTCGTGCTCACCTGGATCCCACTGTTGCGACGTTTCGGGCGCATGTTGTGCGTGTAT 173131
QY 661 GCAAAACAATTCATCCCGGGTACATATTTCAATTTTGAAGTTCATGCTGTGTGTGATCCG 720
DB 173130 GGAGAACAAACAGACGTCGCCGGGACCTTCTTCATCCTCAGCTTCATGTTGTGGGTGGTGG 173071
QY 721 ATCGGATGTGGCGGGTATATCGCGGTGTGTTTCTTGTGATCGACCCCAATGGCGCGCTT 780
DB 173070 GTCCGATGTGGGTGGTATGACAGTGGCGGTGTTCTTCGGGCGCCACCCATGGCGCGGC 173011
QY 781 GTGAGTCCGAAGAAGCTCTTGGGAAGGCTTTCGCGGCTCCATGTGTTTAGGATCGGTAC 840
DB 173010 GGTGAGCCGAGAAATCTCGGAGGGTTTTCGGGTTCCATCTCTCGGGATCGGTAC 172951
QY 841 TGGTCACTCAGTGTTCATCTCTGCTCGATCACCACTGAGTGGGTGTGATCTTTGG 900
DB 172950 CGGTGCCATCAGGTGCATTTCTCTGCTGCACCACTGGTGTGCTCGGTGTCTATCTCTGG 172891
```

QY 901 TTGTGCGCCCTAGTTGCTGGCGCACGCTGGTGACTTGGTGTGAGTCGCGAGTTCAAACGGGA 960
 Db 172890 TATCTGCTGCTGTTGCTGGCGCACCCCTGGTGACCTGGTGGATCCCAATCCCAACGGGA 172831
 QY 961 TTTGGGCAATCAAGGATATGTCGAACCTCTTCCAGGCCACCGCGGATGATGACCGGTTT 1020
 Db 172830 CCTCGGCATCAAGGACATGTCCTCAATCTGCTGCCCGGACACGGTGGCTGATGGACCGTCT 172771
 QY 1021 GGATGCACTGCTCCCGCGCGCATGGTACGCTGCTGATCTGAGTGTGATCAGCAGCTC 1080
 Db 172770 CGACGGGATGCTGCCCTCAGCCATGGTCACTGGTGTGATCTCAGCGTGGTGGAGCTC 172711
 QY 1081 GTATCCGCTGTAAGCTTGGCGGACGCTTTAAGTTCAAAAACTTTGAAAGGCGCTGAGGTG 1140
 Db 172710 CTACGCGGGTGCACGCGAGGTGTACCATCAGGAGGGCGCGTCCCGGGGTAGTAGTT 172651
 QY 1141 CATTAAGGTGTCACTCAGCGCTT-----TTTGGCTGTCAAAGTTTAAAGGCGTTTAC 1194
 Db 172650 CAGCCCGGTGGCGGCCCTCTTTTGAATCCCTGGAGGCTGATCAGCAGCGGCGCTTCT 172591
 QY 1195 GGATTTTCTTAACCTGGCGACCGTACTCAAACTGCGCACGCCAACCAAGCGCCATAA 1254
 Db 172590 GGATCTTCTTGATCTTACGACGGTACTGACATGCGGACCCCGACCGACCATGCA 172531
 QY 1255 TCAATGCAACCGGTATGCTGTCTAGTAGAAACCGATTCGCGGTGG 1300
 Db 172530 TCAACGCAACCGGTATGCTGTCTAGTAGAAACCGATTCGCGGTGG 172485
 RESULT 11
 LOCUS BX248358 348408 bp DNA linear BCT 06-NOV-2003
 DEFINITION Corynebacterium diphtheriae gravis NCTC13129, complete genome;
 segment 5/8.
 ACCESSION BX248358 BX248353
 VERSION BX248358.1 GI:38200215
 KEYWORDS complete genome.
 SOURCE Corynebacterium diphtheriae
 ORGANISM Corynebacterium diphtheriae
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 (bases 1 to 348408)
 AUTHORS Cerdeno-Tarraga A.M., Efstathiou A., Dover L.G., Holden M.T.G.,
 Pallen M., Bentley S.D., Besra G.S., Churcher C., James K.D., De
 Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
 Hamlin N., Holroyd S., Jagsis K., Moule S., Quail M.A.,
 Rabinowitz E., Rutherford K., Thomson N.R., Unwin L.,
 Whitehead S. and Barrell B.G. Parkhill J.
 TITLE The complete genome sequence and analysis of Corynebacterium
 diphtheriae NCTC13129
 JOURNAL Nucleic Acids Res. 31 (22), 6516-6523 (2003)
 PUBMED 14602910
 REFERENCE 2 (bases 1 to 348408)
 AUTHORS Cerdeno-Tarraga A.M.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-2003) Cerdeno-Tarraga A.M., submitted on behalf
 of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust
 Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:
 amct@sanger.ac.uk
 FEATURES
 Location/Qualifiers
 1..348408
 /organism="Corynebacterium diphtheriae"
 /mol_type="genomic DNA"
 /strain="NCTC13129"
 /db_xref="taxon:1717"
 /note="biotype gravis"
 complement(58..612)
 /gene="apt"
 /locus_tag="DIP1369"
 complement(58..612)
 /gene="apt"
 /locus_tag="DIP1369"
 /EC_number="2.4.2.7"

/note="Similar to Corynebacterium glutamicum adenine
 phosphoribosyltransferase Apt SWALL:APT CORGL
 (SWALL:O87330) (185 aa) fasta scores: E(): 4.9e-42, 66.27%
 id in 172 aa, and to Escherichia coli adenine
 phosphoribosyltransferase Apt or B0469 SWALL:APT ECOLI
 (SWALL:P07672) (183 aa) fasta scores: E(): 5.7e-19, 40.85%
 id in 164 aa"
 /codon_start=1
 /transl_table=11
 /product="adenine phosphoribosyltransferase"
 /protein_id="CAE49900.1"
 /db_xref="GI:38200216"
 /translation="MTVPLYADAREAIACKTFVDFPVPGVIFEDLPVLADAAAFS
 LVELAKNALELPAEGLEDEKVLVDVLTATGGTLDAAKLEACGATVSGYAVVLE
 VDLGGRRLNDAPLVINESATA"
 complement(61..531)
 /gene="apt"
 /locus_tag="DIP1369"
 /note="HMPfam hit to PF00156, Phosphoribosyl transferase
 domain"
 complement(202..240)
 /gene="apt"
 /locus_tag="DIP1369"
 /note="ScanRegExp hit to PS00103, Purine/pyrimidine
 phosphoribosyl transferases signature."
 complement(668..2419)
 /locus_tag="DIP1370"
 complement(668..2419)
 /locus_tag="DIP1370"
 /note="Similar to Corynebacterium glutamicum
 dipeptide-binding protein DciA SWALL:O87329
 (EMBL:AF038651) (502 aa) fasta scores: E(): 4e-83, 48.61%
 id in 434 aa, and to Mycobacterium tuberculosis
 hypothetical lipoprotein Rv2585C precursor or MT2662 or
 MTCY227.16 SWALL:YP85 MYCTU (SWALL:Q50636) (557 aa) fasta
 scores: E(): 7.1e-30, 29.09% id in 574 aa"
 /codon_start=1
 /transl_table=11
 /product="Conserved hypothetical protein"
 /protein_id="CAE49901.1"
 /db_xref="GI:38200217"
 /translation="MCGQLDRVDSILVGMVSYNPPSPQIGNVLVGKAKOTVIRIP
 STNLKEVVRVCGLLSVMAVALSSCAHPESSLSKSNQLYAVNDLVTNAGSTE
 GASTNAQLAGLYPAPVSGPKGLINLSKASQVLPGLNKILVITISDAVSQGL
 EATCDXLLATAGTYDFVDSYNPLMKQIDKECKPRDREVTVTKENFGPRWQL
 FPAGVILPAHAIRKXGVTLEDNTVLQSDVDALQPIAEVNRGFSLDQDPDLQVS
 SGPKYKRVGRGEVYLTRNPHYGQPAEIESVWVPKGNLKEIKDTGSLIEAVTS
 SKEDWLDNDSTNSFDVQAGSGVLTELLANAGVFATRENQAFASCIQSSIASE
 SSELGSVAVPTARTIRANDPMAGHVXDLTEAHLHSDVEOARLQAGOTIAGVEGPD
 ARMEALVANIRSCFAGIIVDAHQESGIRDLSRFTTHGQYQFDKIGIDIVLKA
 LDPATFESVATTSNNIAATQAEULSTWDVFSIFLASQPRVVFVDRVRNVNLTDL
 AGIWNMDRWFPSKSK"
 complement(2421..2473)
 /note="Score 66: 22/22 (100%) matches, 0 gaps"
 complement(2489..3601)
 /gene="secp"
 /locus_tag="DIP1371"
 complement(2489..3601)
 /gene="secp"
 /locus_tag="DIP1371"
 /note="Similar to Corynebacterium glutamicum SecF
 SWALL:Q9AE06 (EMBL:AF038651) (403 aa) fasta scores: E():
 4.4e-84, 62.14% id in 383 aa, and to Escherichia coli
 protein-export membrane protein SecF or B0409 or Z0508 or
 ECS0460 SWALL:SECF ECOLI (SWALL:P19674) (323 aa) fasta
 scores: E(): 1.7e-11, 25.14% id in 338 aa"
 /codon_start=1
 /transl_table=11
 /product="Putative protein export membrane protein"
 /protein_id="CAE49902.1"
 /db_xref="GI:38200218"
 /translation="MSGNSVLSKLYTGDDGLDFVGRRTWYRTTLLIIVATCFIAIVR

```
misc_feature
/feature_id="G1:38200220"
/db_xref="G1:38200220"
/translation="MAKRIPLNLDNFVKCSVSDSASVAYHSQBKLLFDRVHFRIDGK
QREADGGVLLDHLHRS"
gene
complement (5611..5904)
/locus_tag="Dip1374"
CDS
complement (5611..5904)
/locus_tag="Dip1374"
/locus_tag="Dip1374"
/notes="Similar to Corynebacterium glutamicum hypothetical
13.5 kDa protein SWALL:Q9AB08 (EMBL:AF038651) (121 aa)
fasta scores: E(): 7.2e-10, 34.8% id in 89 aa, and to
Mycobacterium tuberculosis hypothetical 13.0 kDa protein
Query Match 26.3%; Score 342.2; DB 1; Length 348408;
Best Local Similarity 63.1%; Pred. No. 5.5e-95;
Matches 527; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
QY 237 TGCCCAAAACCCAAAATAATGCGGTCGAGATCTCAAAAGCTGCCATTCCTGTTGGGATCG 296
DB 142124 TGCCCTAAACCGAAGAAATTCGGCTGTCGAAACCTGAAAGCCGCGATCAGCGTTGAATTG 142065
QY 297 GACTGGGGTCTCTGTTCTTTTGGGATGTCCTTAAGCCCATGGGTTGGTACATCTCG 356
DB 142064 GCTTGGCGCCCTCGTATTTGCTCGCTATCTTTGTTATTCCTTTTGGTTGGTATTCATTAG 142005
QY 357 TTGCAAGTTTATGCTGCAGCAACATGGAAGTTGGTAGCAGACTTAAAGAAGCGCGCT 416
DB 142004 TGGCTATCCCATCGAGTGGCAATTTGGAGGTAGAGCGCGTTTAAAGAAGCGGTT 141945
QY 417 ATCAATTCGCACTCGCGATATGATCATCGGCGGTGAGCAATCATCTGGCTGTCTATGCC 476
DB 141944 ATCTTTGCAAGTTGGGTCTGCTTATTTGGCGGCAAACTATCTGCGGTTAGTTGGC 141885
QY 477 CATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGCCACTGCTGTTGGTGTCTGATGTA 536
DB 141884 CATTTGGTCCGAAGGATAGTGGCTGTTTCTGTCGGCTGTTACTCGTACGATGTTG 141825
QY 537 TCCGAATTTTCAATAGGACGCGAAAGAAAGACCGCGCAACTATTTAGAGGACACCTCTG 596
DB 141824 GTCCGCTTTTCCATCATGCTGCTTCTATGCCACCGAAACTATCTGCGAGATAC 141765
QY 597 TGGGCATCTTGTGTCACCTGATCCATCTTCCGAGCTTCGTCGAGTCTGCTGTCG 656
DB 141764 TCGCAATTTTGTGTTTACGTGATCCCACTTTTGGTAGTTTGGCGCAATGCTTTCG 141705
QY 657 TGATGCAAAACAATTCATCCCGGGTACATATTTTCAATTTTGAAGTTTCATGCTGTGTGA 716
DB 141704 TGTTTGAACAGAAACAGACACCGGAAAGTACTTCAATGTCACATTTGTTGGTTA 141645
QY 717 TCGCATCGGATGTGGCGGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGCGCG 776
DB 141644 TCGCCTCTGACGTAGGCGGTATATTTGCGAGCGGTGATGTTGCGATCGCACCGAC 141585
QY 777 CGTTGGTGTCTCGAAGAGTCTTTGGAAGGCTTTCCCGGCTCCATGTTCTTAGGATCGG 836
DB 141584 CTGCGGTAGCCCCAAGAAATCTGGAAGGCTTTATCGGTTCTGTAGTCTTGGCATGA 141525
QY 837 TCACTGGTGCACCTCAGTGTTCATCTCCTGTCGATCACCACCTGTTGGATGGGTGATCT 896
DB 141524 TTGTTGGAGCATTCATCTGTGCTTATTTATAGGCGATCAATGTTGGTGGGATGATTC 141465
QY 897 TGGGTTGTGCCCTAGTGTGTGGCGCACCGTGGTGACTTGGTTGAGTGCAGTTCGAAAC 956
DB 141464 TCGGATTTGGACTGTGTGATCTGCGGACACTTGGCGGATCTTGTGCAATCCAGTTTGA 141405
QY 957 GCGATTTGGGCATCAAGGATATCTCGAACTTCTCCAGGCCACGCGGATGATGAC 1016
DB 141404 GGGAACTCGGATCAAGACATCTCTGCGATTTTGGCGGCAATGTTGGTCTGATGATC 141345
QY 1017 GTTTGAGTGCATCTCTCCCGGCGCGATGTTGAGTGGTGTGATCTCCTGAGTGTGAT 1071
DB 141344 GGCTTGACGTTATGTTGCCATCGGCTATGTTGATGATGTTGAGTGTGAT 141290
```

```
misc_feature
/feature_id="G1:38200220"
/db_xref="G1:38200220"
/translation="MAKRIPLNLDNFVKCSVSDSASVAYHSQBKLLFDRVHFRIDGK
QREADGGVLLDHLHRS"
gene
complement (5611..5904)
/locus_tag="Dip1374"
CDS
complement (5611..5904)
/locus_tag="Dip1374"
/locus_tag="Dip1374"
/notes="Similar to Corynebacterium glutamicum hypothetical
13.5 kDa protein SWALL:Q9AB08 (EMBL:AF038651) (121 aa)
fasta scores: E(): 7.2e-10, 34.8% id in 89 aa, and to
Mycobacterium tuberculosis hypothetical 13.0 kDa protein
Query Match 26.3%; Score 342.2; DB 1; Length 348408;
Best Local Similarity 63.1%; Pred. No. 5.5e-95;
Matches 527; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
QY 237 TGCCCAAAACCCAAAATAATGCGGTCGAGATCTCAAAAGCTGCCATTCCTGTTGGGATCG 296
DB 142124 TGCCCTAAACCGAAGAAATTCGGCTGTCGAAACCTGAAAGCCGCGATCAGCGTTGAATTG 142065
QY 297 GACTGGGGTCTCTGTTCTTTTGGGATGTCCTTAAGCCCATGGGTTGGTACATCTCG 356
DB 142064 GCTTGGCGCCCTCGTATTTGCTCGCTATCTTTGTTATTCCTTTTGGTTGGTATTCATTAG 142005
QY 357 TTGCAAGTTTATGCTGCAGCAACATGGAAGTTGGTAGCAGACTTAAAGAAGCGCGCT 416
DB 142004 TGGCTATCCCATCGAGTGGCAATTTGGAGGTAGAGCGCGTTTAAAGAAGCGGTT 141945
QY 417 ATCAATTCGCACTCGCGATATGATCATCGGCGGTGAGCAATCATCTGGCTGTCTATGCC 476
DB 141944 ATCTTTGCAAGTTGGGTCTGCTTATTTGGCGGCAAACTATCTGCGGTTAGTTGGC 141885
QY 477 CATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGCCACTGCTGTTGGTGTCTGATGTA 536
DB 141884 CATTTGGTCCGAAGGATAGTGGCTGTTTCTGTCGGCTGTTACTCGTACGATGTTG 141825
QY 537 TCCGAATTTTCAATAGGACGCGAAAGAAAGACCGCGCAACTATTTAGAGGACACCTCTG 596
DB 141824 GTCCGCTTTTCCATCATGCTGCTTCTATGCCACCGAAACTATCTGCGAGATAC 141765
QY 597 TGGGCATCTTGTGTCACCTGATCCATCTTCCGAGCTTCGTCGAGTCTGCTGTCG 656
DB 141764 TCGCAATTTTGTGTTTACGTGATCCCACTTTTGGTAGTTTGGCGCAATGCTTTCG 141705
QY 657 TGATGCAAAACAATTCATCCCGGGTACATATTTTCAATTTTGAAGTTTCATGCTGTGTGA 716
DB 141704 TGTTTGAACAGAAACAGACACCGGAAAGTACTTCAATGTCACATTTGTTGGTTA 141645
QY 717 TCGCATCGGATGTGGCGGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGCGCG 776
DB 141644 TCGCCTCTGACGTAGGCGGTATATTTGCGAGCGGTGATGTTGCGATCGCACCGAC 141585
QY 777 CGTTGGTGTCTCGAAGAGTCTTTGGAAGGCTTTCCCGGCTCCATGTTCTTAGGATCGG 836
DB 141584 CTGCGGTAGCCCCAAGAAATCTGGAAGGCTTTATCGGTTCTGTAGTCTTGGCATGA 141525
QY 837 TCACTGGTGCACCTCAGTGTTCATCTCCTGTCGATCACCACCTGTTGGATGGGTGATCT 896
DB 141524 TTGTTGGAGCATTCATCTGTGCTTATTTATAGGCGATCAATGTTGGTGGGATGATTC 141465
QY 897 TGGGTTGTGCCCTAGTGTGTGGCGCACCGTGGTGACTTGGTTGAGTGCAGTTCGAAAC 956
DB 141464 TCGGATTTGGACTGTGTGATCTGCGGACACTTGGCGGATCTTGTGCAATCCAGTTTGA 141405
QY 957 GCGATTTGGGCATCAAGGATATCTCGAACTTCTCCAGGCCACGCGGATGATGAC 1016
DB 141404 GGGAACTCGGATCAAGACATCTCTGCGATTTTGGCGGCAATGTTGGTCTGATGATC 141345
QY 1017 GTTTGAGTGCATCTCTCCCGGCGCGATGTTGAGTGGTGTGATCTCCTGAGTGTGAT 1071
DB 141344 GGCTTGACGTTATGTTGCCATCGGCTATGTTGATGATGTTGAGTGTGAT 141290
```



```

YVGYGLYELRLIAGVANGPQDAVIAAGRLGALAGWYNQHGAWFVALLVVLVVG
FAGTWFRVR"
/translation="5731..6264"
/gene="MT2946"
/contig="us-09-853-641-1"
/length="6264"
/notes="similar to SP:Q10804 PID:1403398
PID:1707631 GB:AL123456; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="soluble secreted antigen MPT53 precursor"
/protein_id="AAK47271.1"
/db_xref="GI:13882722"
/translation="MRSVMSRLVSPKADGIVAJAVLVMFLGANTPRAVADE
RLGFTAILSGAFDGSLOKFAVLNFWPFCNAPSLSOVAANPAVTFVGI
ATRADVGMSQSVSKYNLNTNDADGVIWIRNFWQPAFYFADGTSFVNPT
AAMSQELSGRVAALTS"
/complement(6438..7532)
/gene="MT2947"
/complement(6438..7532)
/gene="MT2947"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK47272.1"
/db_xref="GI:13882723"
/translation="MVPELMPDEPRGPRRHLDLDAAGPASAVALGLPAPRAKOL
AHQYGLRIADPROMDLPAAVRDRIAGAFNLLTASADITCDAGOTKRLAVDNG
TMFESVLMRYPRNTVCISQAGCGACFPFCATCGGGLTRNLSTAELEQVRAAAL
RDFGRLSNVFMGEPGLANVLAQVRIARFPGSGISARAATVSTVGLAPA
IRNLADARLGVTLALSDGLRDLTPVNNRWIRSEALDAARYANVTGRRVIE
YALRDYNDOPRALDLGLRHLRVGLPLAHVLIPLNPTPGSDWDASPKVEFEVFR
VRAGVSVCTVRDTEGREISACGLAAVGG"
/complement(7555..8604)
/gene="MT2948"
/complement(7555..8604)
/gene="MT2948"
/notes="similar to GB:M11330 SP:P06466 PID:1208947
PID:145476 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="phosphatidate cytidyltransferase, putative"
/protein_id="AAK47273.1"
/db_xref="GI:13882724"
/translation="MRSVAOKRISTRPTNTSPKLSWLNTPKASWRSGRSATS
TTNACGNPAEQPARGAKQAPATSPAGSDLEAAIVGLSIGLVLIAVLVFPVPM
VAIVAVATLVATHEVRRRLRAGHLIPVILLGGQAAVLTWPFGVAGLAGGNY
VVCMIWRLFMQDSVTRPTTGGAPSPGNLSDVATVLAHVPLFCFSGAMLVYPNG
SGWVFCMMIAVSDVGVYAVGVLFGRHPMVTTPSPKSWEGFAGSLVCGITATIITA
TFLVKTPIWIGALLGVFLVLTALGLDVLVESQVKRDLGDKMDGRLLPHGGLMDRLGI
LPSAAVAWLVLTLP"
/complement(8498..9055)
/gene="MT2949"
/complement(8498..9055)
Query Match 14.8%; Score 191.8; DB 1; Length 16952;
Best Local Similarity 54.4%; Pred. No. 5.6e-48;
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;
257 GCGGTGAGATCTCAAGCTCCATTCGTGTGGGATCGGACTGGGGTCTCTGTTCTT 316
8388 GCGGCGCTGATTTCGCGCAGCGATCGTGTGGGCTTCATAGGCTAGTCTCATC 8329
317 TTGGGATGTCCTAAGCCATCGGGTGTGATACATCCTCGTTGCGAGTTTATGCTGCA 376
8328 GCGGTGCTGTTCTGTTCCGCGGTTTGGGTGCTCATCGTGGCGCTCGCACCTTGTC 8269
377 GCAACATGGGAATGGTAGCAGACTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436
8268 GCTACCATGAGTGTGCGGAGGTTGCGGAAGCGGCTATCTCATCCCGTTATCCG 8209

```

```

437 ATGATCATCGCGGTTCAGGCAATCATCTGGCTGTATGCGCATTTGGCAGATGGGCATT 496
8208 TTGCTGATTTGGCGGCGAGCGCGGTGTGGCTGACCTTGGCGGTTCGGCGGTTCGGCGCA 8149
497 TTGGCGCTTTTGTGGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
8148 TTGGCGGCTTTGTGTGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8089
557 ACGGAAAAAAGAACCC-----GCAACTATTTTGAGGAGACACC 592
8088 ACGGTGACGCGCCCGACACCGCGGTGCACCGTCCGCGGAAACTACTTTGCGACGTC 8029
593 TCTGTGGGCATCTTCGTGTCTACCTGGAATTCATTGTTCGGAAGCTTCGTCGCGATGCTG 652
8028 TCGGCGACCGGTCTTCTCGCGGTGTGGGTTCCTGCTCTTTTCGGCGCAATGCTG 7969
653 TCGCTGATGCAAAACAATTCATCCCGGATATATTTTTCATTTTTCAGTTCATGCTGTGT 712
7968 GTCTACCCCGAAAAATGGCTCGGATGGG-----TGTTCGATGATGATCGCG 7921
713 GTGATCGCATCGGATGTGGCGGATATATCGCGGTGTGTTCCTTTTGGATTCGCCCAATG 772
7920 GTCACTGCTTCGATGTGGCGGTACGCGGTGGGTGTCTGTTTGGCAAGCATCCGATG 7861
773 GCGCGCTTGTGTGAGTCCGGAAGTCTTGGGAAGCTTTGCCGCTCCATTCCTTAGGA 832
7860 GTTCCGACGATCAGCCCGAAGTCTGTGGGAGGCTTTGCCGCTTGTGTGTGTGCGGG 7801
833 TCGGTCTACTGTGTCTCACTAGTGTTCATCTCTGCTCGATCACCCTGTGTGTGTGTGTGT 892
7800 ATCACCAGCAACCATCATCACCCGAGCTTCTTCTGCGGCAAAACCGCTGGATGTGTGCA 7741
893 ATCTTGGGTGTGCGCTAGTGTGTGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 952
7740 CTGCTCGCGGTGCTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7681
953 AAACGCGATTTGGGATCAAGGATATGTCAACCTTCTTCCAGGCGACCGCGGATGATG 1012
7680 AAACGTGACTCGGATCAAGACATGCGCGCTCTACCGCGGCGACCGCGGTGTGTGTGTGT 7621
1013 GACCGTTTGGATGGCATGCTCCCGCGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1071
7620 GACCGGTGCGAGCATATGCTGCTTCCGCGGTGGCGGCTGTGTGTGTGTGTGTGTGTGTGT 7562
RESULT 13
BX248344/c 318050 bp DNA linear BCT 11-JUN-2003
LOCUS
DEFINITION
11/14
Myobacterium bovis subsp. bovis AF2122/97 complete genome; segment
BX248344 BX248333
VERSION
BX248344.1 GI:131619628
complete genome.
KEYWORDS
Myobacterium bovis subsp. bovis AF2122/97
Myobacterium bovis subsp. bovis AF2122/97
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
REFERENCE
AUTHORS
Garnier,T., Eiglmeyer,K., Camus,J.-C., Medina,N., Mansoor,H.,
Pryor,M., Duthoy,S., Grondin,S., Lacroix,C., Monsemp,C., Simon,S.,
Harris,B., Atkin,R., Doggett,J., Mayes,R., Keating,L.,
Wheeler,P.R., Parkhill,J., Barrell,B.G., Cole,S.T., Gordon,S.V. and
Hawinson,G.
The complete genome sequence of Mycobacterium bovis
Online Publication
PNAS 10.1073/pnas.1130426100 ( Microbiology )
2 (bases 1 to 318050)
Garnier,T.
Direct Submission
Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire
Bacterienne Institut Pasteur 28,rue du Dr Roux 75724 PARIS cedex

```

15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the Mycobacterium bovis sequencing teams, TB Research Group, Veterinary Laboratories Agency Weybridge, Woodham Lane, New Haw, Addlestone, Surrey KT15 3NE, UK. Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK. PTA Annotation, Genopole, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France. Unité de Génétique Moléculaire Bactérienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

FEATURES

Source

1..318050
Location/Qualifiers

/organism="Mycobacterium bovis subsp. bovis AF2122/97"

/mol_type="genomic DNA"

/strain="AF2122/97"

/db_xref="taxon:233413"

complement(99..875)

/locus_tag="Mb2882c"

complement(99..875)

/locus_tag="Mb2882c"

/EC_number="1.1.1.1"

/note="Mb2882c", len: 258 aa. Equivalent to Rv2857c.

len: 258 aa, from Mycobacterium tuberculosis strain H37Rv,

(100.0% identity in 258 aa overlap). Probable short-chain

dehydrogenase/reductase (EC 1.1.1.1), highly similar to

various dehydrogenases e.g. O88068|SCI35.33c PROBABLE

DEHYDROGENASE (SDR FAMILY) from Streptomyces coelicolor

(260 aa), FASTA scores: opt: 1208, E(): 2e-68, (72.35%

identity in 253 aa overlap); Q9I376|PA1649 from

Pseudomonas aeruginosa PROBABLE SHORT-CHAIN DEHYDROGENASE

(253 aa), FASTA scores: opt: 569, E(): 2.1e-28, (39.2%

identity in 255 aa overlap); Q9EX74|MLHA SDR-LIKE ENZYME

from Rhodococcus erythropolis (246 aa), FASTA scores: opt:

567, E(): 2.8e-28, (41.15% identity in 248 aa overlap);

etc. Also similar to many Mycobacterium tuberculosis

dehydrogenases e.g. FABG3|Rv2002|MT2058|MTCV39.16c

PUTATIVE OXIDOREDUCTASE (260 aa), FASTA score: (38.3%

identity in 248 aa overlap). BELONGS TO THE SHORT-CHAIN

DEHYDROGENASES/REDUCTASES (SDR) FAMILY."

/codon_start=1

/transl_table=11

/product="PROBABLE SHORT-CHAIN TYPE

DEHYDROGENASE/REDUCTASE"

/protein_id="CAD96569.1"

/db_xref="GI:31619639"

/translation="WMDSRLGRVAVITGGSGIGLAGRRMRAGATIVGVDDV

EAGGADEUSGLFVFDVDEAVNGLFDGAETGYRIDIAFNAGISPEVDNLIEN

TELAARVQVDVNLKSLVLCRAALRMVLGKSVNTASVFAVMGSAATQSISYAS

KGVLAARVQVQFARQGRVIRVNLALCPGVNTPLQLFALNPERAARMVHVPGLGRF

AEPDEIAAVALFLASDDASITASTFLVDGGISSAVYVPL"

complement(872..2239)

/gene="aldC"

/locus_tag="Mb2883c"

complement(872..2239)

/gene="aldC"

/locus_tag="Mb2883c"

/EC_number="1.2.1.3"

/note="Mb2883c, aldC, len: 455 aa. Equivalent to Rv2858c,

len: 455 aa, from Mycobacterium tuberculosis strain H37Rv,

(100.0% identity in 455 aa overlap). Probable aldC,

aldehyde dehydrogenase (EC 1.2.1.3), similar to many e.g.

O88069|SCI35.34c PUTATIVE ALDEHYDE DEHYDROGENASE from

Streptomyces coelicolor (483 aa), FASTA scores: opt: 1872,

E(): 6.4e-109, (64.5% identity in 448 aa overlap);

Q9FAB1|ALDH|BT-ALDH ALDEHYDE DEHYDROGENASE from Bacillus

thermoleovorans (497 aa), FASTA scores: opt: 1157, E():

2.1e-64, (44.3% identity in 458 aa overlap); O33455|CYMC

P-CUMIC ALDEHYDE DEHYDROGENASE from Pseudomonas putida

(494 aa), FASTA scores: opt: 1149, E(): 6.5e-64, (43.15%

identity in 452 aa overlap);

P40047|DHAS YEAST|ALD5|ALDH5|ALD3|YER073W ALDEHYDE

DEHYDROGENASE from Saccharomyces cerevisiae (Baker's

yeast) (519 aa), FASTA scores: opt: 1091, E(): 2.7e-60,

(38.55% identity in 459 aa overlap);

P80668|FEAB_ECOLI|PADA|MAOB|B1385 PHENYLACETALDEHYDE

DEHYDROGENASE (EC 1.2.1.39) from Escherichia coli strain K12 (499 aa), FASTA scores: opt: 1074, E(): 3e-53, (42.2% identity in 462 aa overlap); etc. Also similar to many M. tuberculosis dehydrogenases e.g. P71823|Rv0768|MTCV369.13 (489 aa), FASTA score: (38.1% identity in 467 aa overlap). Contains PS00687 Aldehyde dehydrogenases glutamic acid active site and PS00070 Aldehyde dehydrogenases cysteine active site. BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY."

/codon_start=1

/transl_table=11

/product="PROBABLE ALDEHYDE DEHYDROGENASE ALDC"

/protein_id="CAD96570.1"

/db_xref="GI:31619630"

/translation="MSTQLINPATEVVLASVDHTDANAVDDAVORASAAQRWARLA

PAQRAAGLRAFAAAQVAHDELAALVANSAGHPVISAWEAGHVDVLAFAASPELTP

SGRQIPVAGGVDTFTNPMVGVITPMNPFMVAIAAPALAAVLAFLVKPABELTP

LTTRMIGELAVAGLEDELLQVPGKTVGERFVTHPDRIKIVFTGSTRVGRVMAG

AAQVQRVTLGLGKSAIVPHDCLEAATAATAPAGVFNAGDCCARILVQRVSIV

DFEMELLEPAVHSIVVGGPGSRATEMGPVSPRAHDKVAGVYDDPAPVAFGTAPGR

GFMEPTVLTPKRGRTVTDTEIFGPNVILFDDEADAISLANDAYGLSGSIWTDLL

SRALRVARAVESGNLSVNSHSFRNTFFGFKQSGVGRLELGPDLAPLQFTETKNVFLA

VGEEM"

complement(2236..3162)

/locus_tag="Mb2884c"

complement(2236..3162)

/locus_tag="Mb2884c"

/EC_number="2.3.5.1"

/EC_number="6.3.5.1"

/note="Mb2884c", len: 308 aa. Equivalent to Rv2859c,

len: 308 aa, from Mycobacterium tuberculosis strain H37Rv,

(100.0% identity in 308 aa overlap). Possible

amidotransferase (EC 6.3.5.1 or 2.3.5.1) equivalent (but

longer 58 aa) to Q9CEU9|ML1573 POSSIBLE AMIDOTRANSFERASE

from Mycobacterium leprae (249 aa), FASTA scores: opt:

1226, E(): 3e-64, (71.55% identity in 239 aa overlap).

Also similar to other amidotransferases and hypothetical

proteins, but shorter in N-terminus e.g. O88072|SCI35.37

HYPOTHETICAL 25.3 KDA PROTEIN from Streptomyces coelicolor

(242 aa), FASTA scores: opt: 683, E(): 1.2e-32, (47.65%

identity in 235 aa overlap); AAK79730|Q97188|CAC1764

PREDICTED GLUTAMINE AMIDOTRANSFERASE from Clostridium

acetobutylicum (241 aa), FASTA scores: opt: 458, E():

1.6e-19, (32.95% identity in 246 aa overlap);

AAK75201|Q97QW9|SP1089 GLUTAMINE AMIDOTRANSFERASE CLASS I

from Streptococcus pneumoniae (229 aa), FASTA scores: opt:

431, E(): 5.6e-18, (34.75% identity in 236 aa overlap);

etc. Contains three 17 aa repeats at the N-terminus very

similar to those in other Mycobacterium tuberculosis

proteins e.g. Q10699|Y130 MYCTU|Rv2090|MT2151|MTCV49.30

PUTATIVE 5'-3' EXONUCLEASE Rv2090 (EC 3.1.11.1)."

/codon_start=1

/transl_table=11

/product="POSSIBLE AMIDOTRANSFERASE"

/protein_id="CAD96571.1"

/db_xref="GI:31619631"

/translation="MDLSASRSDGGDPLRPASPLRSPVSDGGDPLRPASPLRSPVS

DGGDPLRPASPLRSPVGLRSPVGLTAYLEQVTRGVMDIPAGVLPAYFEGITMAGS

VAVLLPQVPDPSPGVGLSDHALVITGGYDLDPAAAYGQEPHPATDPPRFGDWEF

ALLRGLQRGMPVLGICRGTVGNALGVGLDILGHLPLDILGSHRACNGVTRPLPVHT

ASGTRLAELIGESADVPCYHQAIDVQVBGLVNSAVDVGVIETALELPDGTFLVAVQW

HPEKSLDDRLFKALVDAASGYAGRQQAEP"

complement(3143..4516)

/gene="glnA4"

/locus_tag="Mb2885c"

complement(3143..4516)

/gene="glnA4"

/locus_tag="Mb2885c"

/EC_number="6.3.1.2"

/note="Mb2885c, glnA4, len: 457 aa. Equivalent to Rv2860c,

len: 457 aa, from Mycobacterium tuberculosis strain H37Rv,

(100.0% identity in 457 aa overlap). Probable glnA4,

glutamine synthetase class II (EC 6.3.1.2), similar to


```
CC (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/)
XX Key Location/Qualifiers
FH source 1. .348676
FT /db_xref="taxon:83332"
FT /mol_type="genomic DNA"
FT /organism="Mycobacterium tuberculosis H37Rv"
FT /strain="H37Rv"
FT complement(148..183)
FT repeat_unit /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_region complement(148..4539)
FT repeat_unit /note="4392 bp direct repeat region"
FT complement(222..257)
FT /note="36 bp direct repeat, 35 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(298..333)
FT /note="36 bp direct repeat, 35 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(374..409)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(447..482)
FT /note="36 bp direct repeat, 35 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(519..554)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(590..625)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(664..699)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(740..775)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(811..846)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(884..919)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(958..993)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(1031..1066)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(1104..1139)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(1176..1211)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(1248..1283)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(1322..1357)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(1396..1431)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(1467..1486)
FT /note="20 bp partial direct repeat, CCGCGAGGGGACGGAAC,
FT of sequence GTGCTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_region complement(1486..2860)
FT /note="IS6110-11. len: 1375 bp. Insertion sequence IS6110."
FT /insertion_seq="IS6110-11"
FT CDS complement(1529..2467)
FT /note="RV2814c, (MTCY16B7.29), len: 312 aa. Probable
FT transposase, highly similar to others e.g.
FT P97137|RV0796|MIV042.06 PUTATIVE TRANSPOSASE FOR INSERTION
FT SEQUENCE ELEMENT IS986/IS6110 from Mycobacterium
FT tuberculosis (328 aa), FASTA scores: opt: 2103, E():
FT 6.1e-132, (100.0% identity in 312 aa overlap); etc. Start
FT unlikely."
FT /transl_table=11
FT /function="REQUIRED FOR THE TRANSPOSITION OF THE INSERTION
FT ELEMENT IS6110."
FT /locus_tag="RV2814c"
FT /product="PROBABLE TRANSPOSASE"
FT /protein_id="CAB03675.1"
FT /translation="LITRFADHOGHREGPCDGLRWGVESICTOLTELGVPIAPSTVYDH
FT INRESRRELRLDGLKEHLSRVHAAHYGVGARKVLTINREGIVATVETLRLTKLG
FT LSGTRGKARRTTIADPATARPADLVQRRFPGPAPNRLWADLVTVSWAGPAYVAVFT
FT DAYARILGWVASTWATSMVLDLALEQAIWTQQEGVLDLKDVIHHTDORGSTYIRFS
FT ERLAAGIQPSVNGATSSYDNLAEITINGLYKTELKPKGKWRSTIEDVELATRVWDF
FT NHRLIYQCGDVPVPELEAAAYVQRRPAAG"
FT complement(2464..2790)
FT /note="RV2815c, (MTCY16B7.28), len: 108 aa. Probable
FT transposase, identical from aa 51 with P19772|YIA2_MYCTU
FT PUTATIVE TRANSPOSASE (INSERTION ELEMENT IS986) from
FT Mycobacterium tuberculosis (59 aa), FASTA scores: opt: 365,
FT E(): 1.1e-19, (96.6% identity in 59 aa overlap); and other
FT transposases."
FT /transl_table=11
FT /function="REQUIRED FOR THE TRANSPOSITION OF THE INSERTION
FT ELEMENT IS6110."
FT /locus_tag="RV2815c"
FT /product="PROBABLE TRANSPOSASE"
FT /protein_id="CAB03633.1"
FT /translation="MSGSSRRYPPELRERVRVMAEIRGHDSWAAISEVRLIVGV
FT CAETVRKVRQAVDAGARPPTTSEAEKRLREDNAELRRANAILKTASAFPAEILD
FT RPAR"
FT complement(2845..2860)
FT /note="18 bp partial direct repeat, GTCGTCAGACCCCAAA, of
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(2901..2936)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(2976..3011)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3049..3084)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3121..3156)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3193..3228)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3266..3301)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3338..3373)
FT /note="36 bp direct repeat, 32 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3399..3434)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3476..3511)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3548..3583)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3624..3659)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT repeat_unit complement(3701..3736)
FT /note="36 bp direct repeat, 36 out of 36 bp identical to
FT sequence GTCGTCAGACCCCAAAACCCCGAGGGGACGGAAC"
FT
```



```

FT repeat_unit complement (3774..3809)
Query Match 14.8%; Score 191.8; DB 15; Length 348676;
Best Local Similarity 54.4%; Pred. No. 1.2e-47;
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;
Qy 257 GCGGGTCGAGATCCAAAGTCGCTGTGGGATCGGACTGGGGGTCCTGGTTCCT 316
Db 72497 GCGGCGGTGATTTGCGCGAGCGATCGTGTGGGCTTCCCATAGGCTAGTCTCATC 72438
Qy 317 TTGGGGATTCCTAAGCCCAAGGGTGTGATCATCTCGTTCAGAGTTTATGGCTGCA 376
Db 72437 GCGGTGTGTGTTTCGTTCCGCGGTTTGGGTTCCCATCGTGGCGTCGCGACCTGGTC 72378
Qy 377 GCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436
Db 72377 GCTACCATGAGTGTGGGAGGTTGCGGAAGCGGCTATCTCATCCGGTATCCCG 72318
Qy 437 ATGATCATCGCGGTTCAGGCATCATCTGGCTGTCTATGGCCATTTGGCAGATGGGCAAT 496
Db 72317 TTGCTGATTTGGCGGCAGCGCGGTGTGGCTGACCTGGCGGTTTCGGCGGTCGGCGCA 72258
Qy 497 TTGGCGCTTTTGTGGCCACTGTGTGGTGTGATGATTTCCGAATTTCTCAATGGC 556
Db 72257 TTGGCGGCTTTGTGGCATGCTGTGTGTGCAATGATTTGGGACTGTTTCAGGAC 72198
Qy 557 ACGGAAAAGAGGCC-----GCAACTATTTAGGGACACC 592
Db 72197 AGCGTGACGCGCCGACGACCGCGGTCACCGTCGCGGGAACCTACTTGTCCGAGTC 72138
Qy 593 TCTGTGGCATCTTGTGCTCACTGATTCATGATTCATGTTCCGAAGCTTCGTCGATGCTG 652
Db 72137 TCGGCCACGCTCTTCTGCGCGGTGTGGTCCATTTGCTGCTCTTTCGGCGCAATGCTG 72078
Qy 653 TCGCTGATGCAAAACAATTCCATCCCGGTACATATTTGACGTTTCATGCTGTGT 712
Db 72077 GTCATCCCGGAAATGGCTCGGATGG-----TGTTCTGCATGATCGCG 72030
Qy 713 GTGATCGCATCGGATGCGGGGTATATCCGGGTGTGTTCTTTGGATCGACCCCAATG 772
Db 72029 GTCATCGCTTCGATGTCGGCGGTACCGCGGTGTGTTTGGCAAGCATCCGATG 71970
Qy 773 GCGCGTGTGTGAGTCGCAAGACTCTTGGAGGCTTTGCGGCTCCATTTCTTAGGA 832
Db 71969 GTTCCGACGATCAGCCGGAAGTGTGGAGGCTTTGCGGTTTCTGTTGTGCGG 71910
Qy 833 TCGETCACTGTGCACCTCAGTGTTCCTCTCTGATCACCACTGCTGTGAGATGGTGTG 892
Db 71909 ATCACCAGCAAGATCATCACCAGCACTTCTCTGTGCGCAAAACGCGTGGATTTGCA 71850
Qy 893 ATCTTGGGTGTGCCCTAGTTGTGTGCCACGTTGGTGTGCTTGGTGTGATCGCAGTTC 952
Db 71849 CTGCTCGCGGTGCTTTTCTGTGCTCACCACCGCTGGGCGACCTGGTGGATCGCAGTC 71790
Qy 953 AAACCGGATTTGGGCATCAAGGATATGTCGAACCTTTCTCCAGGCGCACGCGGATTTGATG 1012
Db 71789 AAACGTGACCTCGGATCAAGACATGGCGCGCTGCTACCGGCGACGCGGCTCATG 71730
Qy 1013 GACGTTTGGATGGAATGCTCCCGCGCGGATGATGATGCTGTTGATCTCTGAGTGTGAT 1071
Db 71729 GACCGGCTCGAGGCATAGTCTGCTTCCGCGGTGGCGGCTGGATAGTCTCACACTGCT 71671

RESULT 15
MSGB1529CS
LOCUS 36985 bp DNA linear BCT 17-DEC-2001
DEFINITION Mycobacterium leprae cosmid B1529 DNA sequence.
ACCESSION L78824
VERSION L78824.1 GI:1377769
KEYWORDS Mycobacterium leprae
SOURCE Mycobacterium leprae
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

```

```

Corynebacterineae; Mycobacteriaceae; Mycobacterium.
1 (sites)
Eiglmeier, K., Honore, N., Woods, S.A., Caudron, B. and Cole, S.T.
Use of an ordered cosmid library to deduce the genomic organization
of Mycobacterium leprae
Mol. Microbiol. 7 (2), 197-206 (1993)
93188700
PUBMED 846027
REFERENCE
2 (bases 1 to 36985)
Smith, D.R., Richerich, P., Rubenfield, M., Rice, P.W., Butler, C.,
Lee, H.M., Kirst, S., Gundersen, K., Abendschan, K., Xu, Q., Chung, M.,
Deloughery, C., Aldredge, T., Maher, J., Lundstrom, R., Tulig, C.,
Falls, K., Imrich, J., Torrey, D., Engelstein, M., Breton, G., Madan, D.,
Nietupski, R., Seitz, B., Mao, J.I. et al.
Multiplex sequencing of 1.5 Mb of the Mycobacterium leprae genome
Genome Res. 7 (8), 802-819 (1997)
97413161
PUBMED 9267804
ORIGINAL SOURCE TEXT: Mycobacterium leprae (clone: cosmid B1529)
(tissue library: Lorient 6) DNA.
This sequence was generated by the Genome Sequencing Center at
Genome Therapeutics Corporation (Collaborative Research Division),
100
Beaver St., Waltham, MA, 02154. Please contact Doug Smith
(smith@eric.com) for further information. The sequence represents
the
insert of a Lorient 6 cosmid clone from a mapped set of clones
constructed from M. leprae genomic DNA isolated from armadillo
liver
[3]. The sequence may not represent the entire cloned insert of
the
cosmid if an overlapping region was previously sequenced from
another
clone. Coding sequences larger than 60 amino acids were predicted
on
the basis of codon usage and homology information. An attempt was
made
to locate the most probable start site based on codon usage,
homology,
the presence of a Shine-Dalgarno sequence, or overlapping orf that
suggested translational coupling. It is possible that the actual
start
site differs from the one selected.
FEATURES
Location/Qualifiers
1..36985
/organism="Mycobacterium leprae"
/mol_type="genomic DNA"
/specific_host="Dasypus novemcinctus"
/db_xref="taxon:1769"
/clone="cosmid B1529"
/tissue_lib="Lorient 6"
/note="The liver of the armadillo was used to isolate the
Mycobacterium leprae."
Query Match 12.6%; Score 164; DB 1; Length 36985;
Best Local Similarity 52.6%; Pred. No. 3.8e-39;
Matches 440; Conservative 0; Mismatches 375; Indels 21; Gaps 3;
Qy 257 GCGGTCGAGATCTCAAAGTCGCTGTGGGATCGGACTGGGGTCCCTGGTTCCT 316
Db 8360 GCGGGACGTAATCGCCCGCGATCGCGGTAAAGTATTTGGTGTCTCGTC 8419
Qy 317 TTGGGATTTCTCTAAGCCCATGGGTTGGTATCATCTCTGTCAGCTTTTATGGCTGCA 376
Db 8420 GCAACGCTGTGTTTGTCTCGGGAATCTGGTGTCTTGTGCGCTGCGCCATTTTCGTT 8479
Qy 377 GCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436
Db 8480 GCTAGCCATGAGGTGTGGCGGCTACGGAAGCTGGATATGTGATTCGGGTATCCCG 8539
Qy 437 ATGATCATCGCGGTGAGCAATCATCTGCTCATGCGCATTTGCGACGATGGCAATT 496

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2004, 13:28:22 ; Search time 53 Seconds
(without alignments)
1583.332 Million cell updates/sec

Title: US-09-853-641-2
Perfect score: 1563
Sequence: 1 MNEPEQHRSRMKPKXNA.....LPAMVTWLLISVSSYPS 297

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1563	100.0	297	5	AAG80226 C. glutam
2	1495	95.6	285	4	AAB76559 Coryneb
3	1495	95.6	285	4	AAB80192 Coryneb
4	1495	95.6	285	4	AAG91962 C glutami
5	1064	68.1	292	6	ABU26046 Protein e
6	725.5	46.4	311	6	ABU34252 Protein e
7	719	46.0	306	6	ABU34659 Protein e
8	719	46.0	306	6	ABU36865 Protein e
9	714	45.7	312	6	ABU35960 Protein e
10	498.5	31.9	305	6	ABM65084 Propionib
11	428	27.4	228	4	AAU65884 Propionib
12	428	27.4	228	6	ARM62403 Propionib
13	350.5	22.4	362	5	ABP66169 Bifidobac
14	310	19.8	271	6	ABU41868 Protein e
15	308	19.7	285	6	ABU28067 Protein e
16	306	19.6	79	5	ABP31480 Human syn
17	303	19.4	273	6	ABU21094 Protein e
18	299	19.1	271	6	ABU39633 Protein e
19	299	19.1	282	6	ABU50039 Protein e
20	298	19.1	280	6	ABU49494 Protein e
21	298	19.1	285	6	ABU48233 Protein e
22	298	19.1	285	6	ABU47356 Protein e
23	295	18.9	263	6	ABU31563 Protein e
24	292.5	18.7	248	6	ABU45081 Protein e
25	291.5	18.7	265	3	AAU75113 Neisseria

ALIGNMENTS

RESULT 1
AAG80226
ID AAG80226 standard; protein; 297 AA.
XX AC AAG80226;
XX AC
DT 22-JAN-2002 (first entry)
XX XX
DE C. glutamicum ATCC 13032 cdsA protein.
XX XX
KW Coryneform bacterium; cdsA; phosphatidate-cytidyltransferase; very slow
KW L-lysine production; amino acid production; medicine; animal nutrition;
KW growth rate.
XX OS Corynebacterium glutamicum.
XX XX
FN DE10021828-Al.
XX XX
PD 08-NOV-2001.
XX XX
PF 04-MAY-2000; 2000DE-01021828.
XX XX
PR 04-MAY-2000; 2000DE-01021828.
XX XX
PA (DEGS) DEGUSSA AG.
XX PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX XX
PI ABU26046 Protein e
XX XX
DR ABU34252 Protein e
XX XX
DR ABU34659 Protein e
XX XX
XX ABU36865 Protein e
XX XX
XX ABM65084 Propionib
XX XX
XX AAU65884 Propionib
XX XX
XX ARM62403 Propionib
XX XX
XX ABP66169 Bifidobac
XX XX
XX ABU41868 Protein e
XX XX
XX ABU28067 Protein e
XX XX
XX ABP31480 Human syn
XX XX
XX ABU21094 Protein e
XX XX
XX ABU39633 Protein e
XX XX
XX ABU50039 Protein e
XX XX
XX ABU49494 Protein e
XX XX
XX ABU48233 Protein e
XX XX
XX ABU47356 Protein e
XX XX
XX ABU31563 Protein e
XX XX
XX ABU45081 Protein e
XX XX
XX AAU75113 Neisseria

full pre mature
mature damaged
very slow

Abp80482 N. gonorr
Abp79682 N. gonorr
Abu37262 Protein e
Abu37670 Protein e
Abu38502 Protein e
Abu33201 Protein e
Abu40735 Protein e
Aay75112 Neisseria
Abu39540 Protein e
Abm67561 Phototrab
Abu30427 Protein e
Abu19701 Protein e
Abu22263 Protein e
Abm8605 Phototrab
Ada35970 Acinetoba
Abu16846 Protein e
Abb92745 Herbicida
Abb92123 Herbicida
Abu43538 Protein e
Abu26650 Protein e

26 288.5 18.5 265 6 ABP80482
27 288.5 18.5 265 6 ABP79682
28 288.5 18.5 265 6 ABU37262
29 286.5 18.3 265 6 ABU37670
30 285 18.2 249 6 ABU38502
31 283 18.1 265 6 ABU33201
32 282.5 18.1 287 6 ABU40735
33 280.5 17.9 265 3 AAY75112
34 280.5 17.9 289 6 ABU39540
35 280 17.9 283 6 ABM67561
36 277 17.7 288 6 ABU30427
37 276 17.7 273 6 ABU19701
38 276 17.7 273 6 ABU22263
39 266.5 17.1 315 6 ABM8605
40 262 16.8 285 6 ADA35970
41 257 16.4 274 6 ABU16846
42 253 16.2 399 5 ABB92745
43 252.5 16.2 430 5 ABB92123
44 249.5 16.0 260 6 ABU43538
45 249 15.9 241 6 ABU26650

CC ATCC 13032 phosphatidate-cytidyl transferase (cdsa) described in the
CC method of the invention
XX
SQ Sequence 297 AA;
Query Match 100.0%; Score 1563; DB 5; Length 297;
Best Local Similarity 100.0%; Pred. No. 8.3e-169;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNEPEQHRSWMPKPKKNAGRDLLKAAITAVGIGLVLLGIVLSPGWYLLVAGFVAAA 60
DB 1 MNEPEQHRSWMPKPKKNAGRDLLKAAITAVGIGLVLLGIVLSPGWYLLVAGFVAAA 60
QY 61 TWEVGSRLKEGYYHLPLPIMIIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFYNGT 120
DB 61 TWEVGSRLKEGYYHLPLPIMIIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFYNGT 120
QY 121 EKEARNYLRDTSVGIIVLTWTLPLFGSFAAMLQNNNSIPGTIYILTFMCLVIAADYGY 180
DB 121 EKEARNYLRDTSVGIIVLTWTLPLFGSFAAMLQNNNSIPGTIYILTFMCLVIAADYGY 180
QY 181 IAGVFFGSHPWAPLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVC 240
DB 181 IAGVFFGSHPWAPLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVC 240
QY 241 ATLGLVDSQKRDIGIKDMNLLPGHGLMDRLDGMPLPAAWVTWLLSVISSYPS 297
DB 241 ATLGLVDSQKRDIGIKDMNLLPGHGLMDRLDGMPLPAAWVTWLLSVISSYPS 297
RESULT 2
AAB76559
ID AAB76559 standard; protein; 285 AA.
AC AAB76559;
DT 11-APR-2001 (first entry)
XX Corynebacterium glutamicum MCT protein SEQ ID NO:100.
XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
XX membrane construction and membrane transport protein; petroleum spill;
XX hydrocarbon degradation; gram positive aerobic bacterium; marker;
XX identification; microorganism; fine chemical production; transformation;
XX genome mapping; genetic engineering.
XX
OS Corynebacterium glutamicum.
XX
XX WO200100805-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000926.
XX
XX 25-JUN-1999; 99US-0141031P.
XX 08-JUL-1999; 99DE-01031454.
XX 08-JUL-1999; 99DE-01031478.
XX 08-JUL-1999; 99DE-01031563.
XX 08-JUL-1999; 99DE-01032122.
XX 09-JUL-1999; 99DE-01032124.
XX 09-JUL-1999; 99DE-01032125.
XX 09-JUL-1999; 99DE-01032128.
XX 09-JUL-1999; 99DE-01032180.
XX 09-JUL-1999; 99DE-01032182.
XX 09-JUL-1999; 99DE-01032190.
XX 09-JUL-1999; 99DE-01032191.
XX 09-JUL-1999; 99DE-01032203.
XX 09-JUL-1999; 99DE-01032212.
XX 09-JUL-1999; 99DE-01032227.
XX 09-JUL-1999; 99DE-01032228.
XX 09-JUL-1999; 99DE-01032228.
XX 09-JUL-1999; 99DE-01032230.
XX 14-JUL-1999; 99DE-01032927.

PR 14-JUL-1999; 99DE-01033005.
PR 14-JUL-1999; 99DE-01033006.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040830.
PR 27-AUG-1999; 99DE-01040831.
PR 27-AUG-1999; 99DE-01040832.
PR 27-AUG-1999; 99DE-01040833.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041395.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042078.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042088.
XX
PA (BADI) BASF AG.
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Habershauer G;
PI
XX WPI; 2001-071486/08.
XX N-PSDB; AAF67792.
XX
PT Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation.
XX
XX Claim 20; Page 324; 1119pp; English.
XX
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAB76510 to
CC AAB76847. The MCT nucleic acids and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention
XX
XX Sequence 285 AA;
SQ
Query Match 95.6%; Score 1495; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 4.2e-161;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 MPKPKNAGRDLLKAAITAVGIGLVLLGIVLSPGWYLLVAGFVAAAATWVGSRLKEGG 72
DB 1 MPKPKNAGRDLLKAAITAVGIGLVLLGIVLSPGWYLLVAGFVAAAATWVGSRLKEGG 60
QY 73 YHLPPLPIMIIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFYNGTEKEARNYLRDTS 132
DB 61 YHLPPLPIMIIIGGQAIILSWPFGTGMILASVATVLLVMYFRIFYNGTEKEARNYLRDTS 120
QY 133 VGIFVLTWTLPLFGSFAAMLQNNNSIPGTIYILTFMCLVIAADYGYTAGVFFGSHPMA 192
DB 121 VGIFVLTWTLPLFGSFAAMLQNNNSIPGTIYILTFMCLVIAADYGYTAGVFFGSHPMA 180
QY 193 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVCATLGLVDSQK 252
DB 181 PLVSPKKSWEFGAGSIVLGSVTGALSVEHFLDHHMMGMVILGCALVVCATLGLVDSQK 240
QY 253 RDLGKIDMGNLLPGHGLMDRLDGMPLPAAWVTWLLSVISSYPS 297
DB 241 RDLGKIDMGNLLPGHGLMDRLDGMPLPAAWVTWLLSVISSYPS 285
RESULT 3
AAB80192
ID AAB80192 standard; protein; 285 AA.
XX

AC AAB80192;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:1118.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
XX fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN W0200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000923.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 01-JUL-1999; 99DE-01030476.
PR 02-JUL-1999; 99US-0142101P.
PR 08-JUL-1999; 99DE-01031415.
PR 08-JUL-1999; 99DE-01031418.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031435.
PR 08-JUL-1999; 99DE-01031443.
PR 08-JUL-1999; 99DE-01031453.
PR 08-JUL-1999; 99DE-01031457.
PR 08-JUL-1999; 99DE-01031465.
PR 08-JUL-1999; 99DE-01031478.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031541.
PR 08-JUL-1999; 99DE-01031573.
PR 08-JUL-1999; 99DE-01031593.
PR 08-JUL-1999; 99DE-01031632.
PR 08-JUL-1999; 99DE-01031634.
PR 08-JUL-1999; 99DE-01031636.
PR 08-JUL-1999; 99DE-01032125.
PR 09-JUL-1999; 99DE-01032126.
PR 09-JUL-1999; 99DE-01032130.
PR 09-JUL-1999; 99DE-01032186.
PR 09-JUL-1999; 99DE-01032206.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032228.
PR 09-JUL-1999; 99DE-01032229.
PR 09-JUL-1999; 99DE-01032230.
PR 14-JUL-1999; 99DE-01032922.
PR 14-JUL-1999; 99DE-01032926.
PR 14-JUL-1999; 99DE-01032928.
PR 14-JUL-1999; 99DE-01033004.
PR 14-JUL-1999; 99DE-01033005.
PR 14-JUL-1999; 99DE-01033006.
PR 12-AUG-1999; 99US-0148613P.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040768.
PR 31-AUG-1999; 99DE-01040832.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041380.
PR 31-AUG-1999; 99DE-01041394.
PR 31-AUG-1999; 99DE-01041396.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042124.
PR 03-SEP-1999; 99DE-01042129.
PR 09-MAR-2000; 2000US-0187970P.
XX
PA (BADI) BASF AG.
XX
XX Pompejus M. Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI
XX WPI: 2001-137957/14.
DR N-PSDB; AAF72311.
XX
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway
PT proteins, useful for producing fine chemicals in microorganisms,
PT including organic acids, nonproteinogenic amino acids, and purine and
PT pyrimidine bases.
XX
PS Claim 20; Page 1676; 1737pp; English.
XX
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
CC nucleic acids are useful for the production of fine chemicals in
CC microorganisms, including organic acids, nonproteinogenic amino acids,
CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
CC vitamins, cofactors, polyketides and enzymes
XX
SQ Sequence 285 AA;
Query Match 95.6%; Score 1495; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 4.2e-161;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 MPKPKNNAGRDILKAAIAVIGIGLVLLGIVLSPGWYLLVAGFMAATWVGSRKKEGG 72
Db 1 MPKPKNNAGRDILKAAIAVIGIGLVLLGIVLSPGWYLLVAGFMAATWVGSRKKEGG 60
QY 73 YHLPPLMIIGGQAIWLSPFGTWGILASFVATVLMVFRIFYNGTEKEARNYLRDTS 132
Db 61 YHLPPLMIIGGQAIWLSPFGTWGILASFVATVLMVFRIFYNGTEKEARNYLRDTS 120
QY 133 VGIFVLTWIPFGSFAAMLQNNNSIPCTYFILTFMLCVIASDVGGYIAGVFFGSHPMA 192
Db 121 VGIFVLTWIPFGSFAAMLQNNNSIPCTYFILTFMLCVIASDVGGYIAGVFFGSHPMA 180
QY 193 PLVSPKKSWEFGAGSIVLGSVTGALSVEFLDDHHMMGVILGCALVVCATIGDLVESQFK 252
Db 181 PLVSPKKSWEFGAGSIVLGSVTGALSVEFLDDHHMMGVILGCALVVCATIGDLVESQFK 240
QY 253 RDLGKIDMSNLLPFGHGLMDRLDGMPLPAAWTVLILSVISSSYPS 297
Db 241 RDLGKIDMSNLLPFGHGLMDRLDGMPLPAAWTVLILSVISSSYPS 285
RESULT 4
AAG91962
ID AAG91962 standard; protein; 285 AA.
XX
XX AAG91962;
AC
XX 26-SEP-2001 (first entry)
DT
XX
XX C glutamicum protein fragment SEQ ID NO: 5716.
DE
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
KW
XX Corynebacterium glutamicum.
OS
XX EP1108790-A2.
PN
XX 20-JUN-2001.
PD

QY 133 VGIFVLTWIPFGSPAMLSIMQNNISIPGTYFILLTFMLCVIASDVGGYAGVFFGSHPA 192
 Db 131 VAIFVLTWIPFGSFAAMLSIFETETAPGYFIVTFMLCVIASDVGGYAGVFFGSHPA 190
 QY 193 PLVSPKSWGFGAGSIVLGSVTCALSVHFLDHRHWMGVILGALVYCATLGLDLVESQPK 252
 Db 191 PAVSPKSWGFGSIVVFGMIVGAFTVSYLLGHQWMMGLVGLFGLVCATLGLDLVESQPK 250
 QY 253 RDLGIKDSMLPCHGGLMDRLDGMIPAAVMTWLISVSS 293
 Db 251 RELGIKDSMAILPGHGLMDRLDGMIPAAVMTWLISVSS 291

RESULT 6

ABU34252

ID ABU34252 standard; protein; 311 AA.

XX AC

ABU34252;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #19779.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Mycobacterium avium.

XX PN WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA38122.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids, required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 62176; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

XX of the nucleic acid inhibits proliferation of a cell. Also included are:

XX (1) a vector comprising a promoter operably linked to the nucleic acid

XX encoding a polypeptide whose expression is inhibited by the antisense

XX nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX polypeptide or its fragment whose expression is inhibited by the

XX antisense nucleic acid; (4) an antibody capable of specifically binding

XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX proliferation or the activity of a gene in an operon required for

XX proliferation; (7) identifying a compound that influences the activity of

XX the gene product or that has an activity against a biological pathway

XX required for proliferation, or that inhibits cellular proliferation; (8)

XX identifying a gene required for cellular proliferation or the biological

XX pathway in which a proliferation-required gene or its gene product lies

XX or a gene on which the test compound that inhibits proliferation of an

XX organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX compound's activity; (11) a culture comprising strains in which the gene

XX product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 311 AA;

Query Match 46.4%; Score 725.5; DB 6; Length 311;

Best Local Similarity 47.5%; Pred. No. 1.6e-73;

Matches 145; Conservative 47; Mismatches 94; Indels 19; Gaps 4;

QY 2 NEPE-----QHRSRMKPKNNAGRDLLAAIAVAGIGLVILGLVSPWGYILVA 54

Db 10 DEPEHAVENTTEGAAGQRAKKTTSRAGRDLLAAIAVAGIGLVILGLVSPWGYILVA 69

QY 55 GFMAATWEGSRLKEGGYHLPLPIMIIIGQAIILSWPFGTMGILAGFVATVILVMYPR 114

Db 70 MAILVASHEVVRRLREAGYIVIPILAGGQUTWLTWTPFHAAGALAGGVTVAFLWR 129

QY 115 IFYNGTEKE-----ARNYLRTDSVGIFVLTWLPFGSFAAMLSLMQNNISIPGYFIL 166

Db 130 LFMQDNKRPPEPPAGSPSANYLRDASATVFLACWVPLFASFAALLVYPADGA--GRVFCL 187

QY 167 TFWLCVIAADVGVIAGVFFGSHPMAPLVSPKSWEGFAGSIVLGSVTCALSVHFLDHH 226

Db 188 --MITVASDVGGYAVGVNLFKGHPVPAISPSPKSWEGLAGSLVIGITATLAATFLAGKA 245

QY 227 WNWGVILGALVVCATLGLDLVESQPKRDLGIDKMSNLLPFGHGLMDRLDGMIPAAVMTWL 286

Db 246 PWVGALLGVVLTCTLGLDLVESQPKRDLGIDKMSNLLPFGHGLMDRLDGMIPAAVMTWL 305

QY 287 ILSVI 291

Db 306 VLTIV 310

RESULT 7

ABU34659

ID ABU34659 standard; protein; 306 AA.

XX AC

ABU34659;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #20186.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Mycobacterium bovis.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

```

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
DR N-PSDB; ACA38529.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 62583; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 306 AA;
XX
Query Match 46.0%; Score 719; DB 6; Length 306;
Best Local Similarity 47.7%; Pred. No. 8.8e-73;
Matches 143; Conservative 46; Mismatches 97; Indels 14; Gaps 3;
QY 2 NEPEQHSEMEM--PKPNAGRDLCRAIAGVIGLGVLLGVLSPWGVILVAGFMAA 59
Db 10 NPAQPPAGAKQQPATETSRAGDRRAIVVGLSGILVLAIVLVFVRVVAIVATV 69
QY 60 ATWEVGSRLKEGGYHLPIMIGQQAIIILSWPFGTWGILASPVATVILVMTFRIPY-- 117
Db 70 ATHEVRLREAGYILVPIPLIGQAAMLTWFFGAVGALAGFGMVVCMWLRFMQD 129
QY 118 -----NCTEARNYLDTSVGFVLTWLPFGSFAAMLNQNNSIPTGYILTFMLC 171
Db 130 SVTRPTTGAPSPGNYLSDVSATVFLAVVPLFCFGAMLVYPEN----GSGWVFCNMIA 185
QY 172 VIASDVGGYIAGVFGGHPMAPLSPKKSNEGFPAGSIVLGSVTVTGALSVEHLDDHWMGV 231
Db 186 VIASDVGGYAGVLFGRHPVMTSPKKSNEGFPAGSLVCGTITATITLVGKTPWIGA 245
QY 232 ILGALVVCATLGLDLVRSQFRDLGIDKMDNLLPGHGLMDRLDGMPLPAAWTVLILSVI 291
Db 246 LLGLVFLVLTALGLDLVESQVRDLGIDKMDRLDGLDGLPESAAVAVITLL 305
XX
RESULT 8
ABU36865
ID ABU36865 standard; protein; 306 AA.

```

Query Match 46.0%; Score 719; DB 6; Length 306;
Best Local Similarity 47.7%; Pred. No. 8.8e-73;

MITCHAM JL, SKEIKY YAW, PERSING DH, BHATIA A, MAISONNEUVE JL;
 ZHANG Y, WANG S, JEN S, LODDES MJ, BENSON DR, JONES R, CARTER D;
 BARTH B, VALLIEVE-DOUGLASS J;
 WPI: 2003-381789/36.
 New Propionibacterium acnes polypeptides and polynucleotides encoding the
 polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 or for stimulating an immune response specific for a P. acnes protein.
 Claim 7; SEQ ID NO 29760; 1481pp; English.
 The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 encoding a Propionibacterium acnes protein. The invention also relates to
 polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 immunogenic fragments of P. acnes polypeptides. The invention
 additionally encompasses expression vectors and host cells comprising a
 polynucleotide of the invention; antibodies against polypeptides of the
 invention; fusion proteins comprising a polypeptide of the invention; a
 method for stimulating an immune response specific for a P. acnes
 polypeptide and an isolated T cell population comprising T cells prepared
 via this method; a vaccine composition (comprising P. acnes polypeptides,
 polynucleotides, antibodies, fusion proteins, T cell populations, or
 antigen-presenting cells that express the polypeptide); a method and kit
 for detecting or determining the presence or absence of P. acnes in a
 patient; and a method for inhibiting the development of P. acnes in a
 patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 proteins, T cell populations or antigen-presenting cells that express the
 polypeptides are useful for diagnosing, preventing or treating acne
 vulgaris, or for stimulating an immune response specific for a P. acnes
 protein. The polynucleotides can also be used as probes or primers for
 nucleic acid hybridisation. The vaccine composition is useful for the
 stimulation of an immune response against P. acnes, or for treating acne,
 and the kit is useful for performing a diagnostic assay. The present
 sequence represents a specifically claimed P. acnes polypeptide which is
 thought to contain an immunogenic region. Note: The sequence data for
 this patent did not form part of the printed specification, but was
 obtained in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences
 Sequence 305 AA:

AAU65884	AAU65884 standard; protein; 228 AA.
ID	AC
XX	AC
XX	AAU65884;
XX	
DT	27-FEB-2002 (first entry)
XX	
DE	Propionibacterium acnes immunogenic protein #26780.
XX	
XX	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW	dermatological; osteopathic; neuroprotectant.
XX	
OS	Propionibacterium acnes.
XX	
XX	W0200181581-A2.
PN	
XX	01-NOV-2001.
PD	
XX	
PF	20-APR-2001; 2001WO-US012865.
XX	
XX	21-APR-2000; 2000US-0199047P.
PR	
PR	02-JUN-2000; 2000US-0208841P.
PR	
PR	07-JUL-2000; 2000US-0216747P.
XX	
XX	(CORI-) CORIXA CORP.
PA	
XX	
XX	Sheiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;
PI	
PI	WPI; 2001-616774/71.
DR	
DR	N-PSDB; AAS9698.
XX	
PT	Propionibacterium acnes polypeptides and nucleic acids useful for
PT	vaccinating against and diagnosing infections, especially useful for
PT	treating acne vulgaris.
XX	
XX	Claim 3; SEQ ID NO 27079; 1069pp; English.
PS	
XX	
XX	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC	polypeptides. The proteins and their associated DNA sequences are used in
CC	the treatment, prevention and diagnosis of medical conditions caused by
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC	pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC	P. acnes is also involved in infections of bone, joints and the central
CC	nervous system, however it is particularly involved in the inflammatory
CC	lesions associated with acne vulgaris. A method for detecting the
CC	presence or absence of P. acnes in a patient comprises contacting a
CC	sample with a binding agent that binds to the proteins of the invention
CC	and determining the amount of bound protein in the sample. The
CC	polypeptides may be used as antigens in the production of antibodies
CC	specific for P. acnes proteins. These antibodies can be used to
CC	downregulate expression and activity of P. acnes polypeptides and
CC	therefore treat P. acnes infections. The antibodies may also be used as
CC	diagnostic agents for determining P. acnes presence, for example, by
CC	enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC	this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 228 AA;

Db 112 DTGGYAVGVLGCKHKLAPRISPKKSWEGFAGSVITAAAFVGNWACLGGLLSAPWAGIVLGV 171

Qy 236 ALVVCATLGDIVESQFKRDGLGDKMSNLLPGHGLMDRLDGLMPLAAMVTLILSVI 291

Db 172 VLALTGTAGDLVESMIKRDAGIKDMSNFLPGHGGVMDRLDVLFSAPFAMVMSLV 227

RESULT 12

ABM62403

ID ABM62403 standard; protein; 228 AA.

XX AC ABM62403;

XX DT 20-OCT-2003 (first entry)

XX DE Propionibacterium acnes membrane-related polypeptide #27079.

XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;

XX KW immunostimulant; immune response; vaccine.

XX OS Propionibacterium acnes.

XX PN WO2003033515-A1.

XX PD 24-APR-2003.

XX PF 11-OCT-2002; 2002WO-US032727.

XX PR 15-OCT-2001; 2001US-00978825.

XX PA (CORI-) CORIXA CORP.

XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX PI Barth B, Vallieve-bouglass J;

XX DR WPI; 2003-381789/36.

XX DR N-PSDB; ACF64627.

XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the

XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

XX PT or for stimulating an immune response specific for a P. acnes protein.

XX PS Claim 3d; SEQ ID NO 27079; 1481pp; English.

XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

XX CC encoding a Propionibacterium acnes protein. The invention also relates to

XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

XX CC immunogenic fragments of P. acnes polypeptides. The invention

XX CC additionally encompasses expression vectors and host cells comprising a

XX CC polynucleotide of the invention; antibodies against polypeptides of the

XX CC invention; fusion proteins comprising a polypeptide of the invention; a

XX CC method for stimulating an immune response specific for a P. acnes

XX CC polypeptide and an isolated T cell population comprising T cells prepared

XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,

XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or

XX CC antigen-presenting cells that express the polypeptide); a method and kit

XX CC for detecting or determining the presence or absence of P. acnes in a

XX CC patient; and a method for inhibiting the development of P. acnes in a

XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

XX CC proteins, T cell populations or antigen-presenting cells that express the

XX CC polypeptides are useful for diagnosing, preventing or treating acne

XX CC vulgaris, or for stimulating an immune response specific for a P. acnes

XX CC protein. The polynucleotides can also be used as probes or primers for

XX CC nucleic acid hybridization. The vaccine composition is useful for the

XX CC stimulation of an immune response against P. acnes, or for treating acne,

XX CC and the kit is useful for performing a diagnostic assay. The present

XX CC sequence represents a specifically claimed polypeptide which is predicted

XX CC to be encoded by an ORF (open reading frame) contained within the P.

XX CC acnes polynucleotides of the invention. Note: The sequence data for this

XX CC patent did not form part of the printed specification, but was obtained

XX CC in electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 228 AA;

Qy Query Match 27.4%; Score 428; DB 6; Length 228;

Best Local Similarity 46.6%; Pred. No. 7.4e-40;

Matches 82; Conservative 32; Mismatches 56; Indels 6; Gaps 3;

Qy 118 NG--TEKEARNYLRDTSVGI FVL TWIPLEFGSPAAMLSLMMQNSIPGT YFILTFLMCLVIAS 175

Db 56 NGVATARGSHGFV KDAASLFTIAYLPLLGCP--VPLMMGDD-CGSRRIATWILSVAS 111

Qy 176 DVGGYIAGVFFGSHPMAPLVSPKKSWEGFAGSVITGALS VHFLLDHHWMMGVILGC 235

Db 112 DTGGYAVGVLGCKHKLAPRISPKKSWEGFAGSVITAAAFVGNWACLGGLLSAPWAGIVLGV 171

Qy 236 ALVVCATLGDIVESQFKRDGLGDKMSNLLPGHGLMDRLDGLMPLAAMVTLILSVI 291

Db 172 VLALTGTAGDLVESMIKRDAGIKDMSNFLPGHGGVMDRLDVLFSAPFAMVMSLV 227

RESULT 13

ABP66169

ID ABP66169 standard; protein; 362 AA.

XX AC ABP66169;

XX DT 19-NOV-2002 (first entry)

XX DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:913.

XX KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;

XX KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;

XX KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;

XX KW rotavirus; food composition; pharmaceutical composition.

XX OS Bifidobacterium longum.

XX PN EP1227152-A1.

XX PD 31-JUL-2002.

XX PF 30-JAN-2001; 2001EP-00102050.

XX PR 30-JAN-2001; 2001EP-00102050.

XX PA (NEST) SOC PROD NESTLE SA.

XX WPI; 2002-668397/72.

XX PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as

XX PT a probe or primer for detecting and/or identifying Bifidobacterium longum

XX PT in a biological sample.

XX PS Claim 3; SEQ ID NO 913; 80pp; English.

XX CC The present invention describes a polynucleotide (I) comprising a

XX CC sequence of a Bifidobacterium genome selected from the nucleotide

XX CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at

XX CC least 90% identity or which hybridises with the sequences given in

XX CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a

XX CC fusion protein, comprising a sequence selected from 1097 sequences given

XX CC in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a

XX CC heterologous polypeptide. (I) has antidiarrheic and antibacterial

XX CC activities, and can be used as an inhibitor of salmonella. (I) (which is

XX CC a probe) is useful for the detection and/or identification of

XX CC Bifidobacterium longum in a biological sample. A carrier containing the

XX CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be

XX CC used for preventing and/or treating diarrhoea brought about by pathogenic

XX CC bacteria and/or rotavirus. The carrier is a food composition selected

XX CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented

XX CC products, ice-creams, fermented cereal based products, milk based

XX CC powders, infant formula, pet food or a pharmaceutical composition

XX CC selected from tablets, liquid bacterial suspensions, dried oral

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 16:56:56 ; Search time 861 Seconds
(without alignments)
7408.351 Million cell updates/sec

Title: US-09-853-641-1

Perfect score: 1300

Sequence: 1 gaagtcgttctgcgaagga.....aggaaaccgattccgctgg 1300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

```

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1300	100.0	1300	12	US-09-853-641-1
2	1300	100.0	3309400	9	US-09-738-626-1
3	978	75.2	978	13	US-10-627-476-99
4	855	65.8	855	9	US-09-738-626-2216
5	342.2	26.3	876	13	US-10-282-122A-17786
6	191.8	14.8	918	13	US-10-282-122A-26399
7	191.8	14.8	921	13	US-10-282-122A-28605
8	168.8	13.0	933	13	US-10-282-122A-25392
9	164	12.6	939	13	US-10-282-122A-27700
10	123.4	9.5	297	11	US-09-864-408A-905
11	117	9.0	258	9	US-09-738-626-2215
12	110.2	8.5	1176	15	US-10-156-761-2610
13	110.2	8.5	9025608	15	US-10-156-761-1
14	69.8	5.4	2256645	17	US-10-470-565-1

15	68.4	5.3	813	13	US-10-282-122A-33608
16	67.2	5.2	813	13	US-10-282-122A-31373
17	60.8	4.7	2731748	17	US-10-297-465A-1
18	59	4.5	843	13	US-10-282-122A-41234
19	58.8	4.5	936	9	US-09-815-243-7776
20	58.8	4.5	936	13	US-10-282-122A-30203
21	58.4	4.5	2017	16	US-10-381-779-95
22	58.2	4.5	789	13	US-10-282-122A-23303
23	58.2	4.5	795	13	US-10-282-122A-24941
24	57.2	4.4	879	16	US-10-369-493-39691
25	57.2	4.4	879	16	US-10-369-493-40043
26	57.2	4.4	927	16	US-10-369-493-39316
27	56.6	4.4	819	13	US-10-282-122A-11441
28	55.8	4.3	855	13	US-10-282-122A-15294
29	55.2	4.2	789	13	US-10-282-122A-24244
30	55.2	4.2	1733	16	US-10-398-221-3425
31	55	4.2	819	13	US-10-282-122A-14003
32	54.6	4.2	748	13	US-10-282-122A-36821
33	54.6	4.2	858	13	US-10-282-122A-39096
34	54.4	4.2	7060	13	US-10-194-163-479
35	53.2	4.1	846	13	US-10-282-122A-26860
36	53.2	4.1	99629	13	US-10-672-787-37
37	53	4.1	858	13	US-10-282-122A-39973
38	52	4.0	792	13	US-10-282-122A-36212
39	51	3.9	795	13	US-10-282-122A-29002
40	50.8	3.9	861	13	US-10-282-122A-32475
41	50	3.8	780	13	US-10-282-122A-12345
42	49.6	3.8	855	13	US-10-282-122A-19807
43	49	3.8	849	13	US-10-282-122A-41779
44	47.8	3.7	798	13	US-10-282-122A-29410
45	47.4	3.6	795	13	US-10-282-122A-38585

ALIGNMENTS

RESULT 1

US-09-853-641-1
; Sequence 1, Application US/09853641
; Publication No. US20040092710A1

; GENERAL INFORMATION:

; APPLICANT: NAMPOOTHIRI, Madhavan

; TITLE OF INVENTION: Nucleotide Sequences Coding for the CdsA Gene

; FILE REFERENCE: 032301 WD 1171

; CURRENT APPLICATION NUMBER: US/09/853,641

; CURRENT FILING DATE: 2001-08-22

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1300

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (200)..(1090)

; OTHER INFORMATION: cdsa gene

; NAME/KEY: RBS

; LOCATION: (187)..(195)

; OTHER INFORMATION:

US-09-853-641-1

Query Match 100.0%; Score 1300; DB 12; Length 1300;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAGTCGTCGCAAGGAAAGGAACGTATGAGGCTCTAGAGACCTTTATCGCAATG 60

Db 1 GAAGTCGTCGTCGCAAGGAAAGGAACGTATGAGGCTCTAGAGACCTTTATCGCAATG 60

Qy 61 GCTTAAGAATTACGTCGTTTCAACGTCGATGGCGGGGAAACGACGCTTTCTTTGCTT 120

Db 61 GCTTAAGAATTACGTCGTTTCAACGTCGATGGCGGGGAAACGACGCTTTCTTTGCTT 120

121 GCAAGAGTGTGTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTCACATTTGTTA 180
121 GCAAGAGTGTGTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTCACATTTGTTA 180
181 GAAATTCAGAGAGAACCCATGAAATGAACCGGAGCAACATCACCGGTCCATGAGATGCC 240
181 GAAATTCAGAGAGAACCCATGAAATGAACCGGAGCAACATCACCGGTCCATGAGATGCC 240
241 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 300
241 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 300
301 GGGGTCTCTGTTCTTTGGGGAATGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360
301 GGGGTCTCTGTTCTTTGGGGAATGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360
361 AGTTTATGCTGCTGAGCAACATGGAAGTGGTAGCAGACTTAAAGAGCGGCTATCA 420
361 AGTTTATGCTGCTGAGCAACATGGAAGTGGTAGCAGACTTAAAGAGCGGCTATCA 420
421 TTGTCACCTGCCATTTATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCCCAT 480
421 TTGTCACCTGCCATTTATGATCATCGCGGTGAGCAATCATCTGGCTGTATGCCCAT 480
481 TGGCAGATGGCATTTGGCGTCTTTTGGCCACTGTTGGTGGTGTATTTCCG 540
481 TGGCAGATGGCATTTGGCGTCTTTTGGCCACTGTTGGTGGTGTATTTCCG 540
541 AATTTTCTACATGCGCAAGAAAGACCGCACTATTGAGGACACCTCTGTGGG 600
541 AATTTTCTACATGCGCAAGAAAGACCGCACTATTGAGGACACCTCTGTGGG 600
601 CATCTCTGCTCAGCTGATTCATTTGTTGGAGCTTCCTGGAGTCTGCTGGATGAT 660
601 CATCTCTGCTCAGCTGATTCATTTGTTGGAGCTTCCTGGAGTCTGCTGGATGAT 660
661 GCAAAACAAATTCATCCCGGGTACATATTTTACCTGATCTGCTGTGATCGC 720
661 GCAAAACAAATTCATCCCGGGTACATATTTTACCTGATCTGCTGTGATCGC 720
721 ATCGATGTCGGCGGTATATCGCGGTGTTGTTGGATCGCAACCAATGGCGCGTT 780
721 ATCGATGTCGGCGGTATATCGCGGTGTTGTTGGATCGCAACCAATGGCGCGTT 780
781 GGTGAGTCGGAAGAGTCTTGGGAAGGCTTTGCGGCTCCATTTGCTTAGGATCGGTAC 840
781 GGTGAGTCGGAAGAGTCTTGGGAAGGCTTTGCGGCTCCATTTGCTTAGGATCGGTAC 840
841 TGGTCACTCAGTGTCTCTCTGCTCGATCACCATCTGGTGGATGGTGTGATCTTTGGG 900
841 TGGTCACTCAGTGTCTCTCTGCTCGATCACCATCTGGTGGATGGTGTGATCTTTGGG 900
901 TTGTCCTCTAGTGTGTCGCAACGTTGGTGAATCTGCTGAGTCGCAATCAACGGGA 960
901 TTGTCCTCTAGTGTGTCGCAACGTTGGTGAATCTGCTGAGTCGCAATCAACGGGA 960
961 TTGTCCTCTAGTGTGTCGCAACGTTGGTGAATCTGCTGAGTCGCAATCAACGGGA 1020
961 TTGTCCTCTAGTGTGTCGCAACGTTGGTGAATCTGCTGAGTCGCAATCAACGGGA 1020
1021 GGATGGATCTCCCGCGCGGATGGTGAATGTTGATCTGAGTGTGATCAGCAGTCTC 1080
1021 GGATGGATCTCCCGCGCGGATGGTGAATGTTGATCTGAGTGTGATCAGCAGTCTC 1080
1081 GTATCCGTCGTAAGCTTGGGCGAGCTTTAAGTTCAAAAACTTTGAAGGCGCTGAGTG 1140
1081 GTATCCGTCGTAAGCTTGGGCGAGCTTTAAGTTCAAAAACTTTGAAGGCGCTGAGTG 1140
1141 CATACGTCGTGCACTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACGATTT 1200
1141 CATACGTCGTGCACTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACGATTT 1200
1201 TTCTTAACTGGCGAGGTACTCAACATGCGCAGCCCAACCAAGCGCCATATCAATG 1260

1201 TTCTTAACTGGCGAGGTACTCAAAATGCGCAGCCCAACCAAGCGCCATATCAATG 1260
1261 CACCGGTATGCTGCTAGTAGGAACCGATTCGGCTGG 1300
1261 CACCGGTATGCTGCTAGTAGGAACCGATTCGGCTGG 1300

RESULT 2
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 100.0%; Score 1300; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTCGTTGCTCGCAAGGAAAGGAACCTGATGAGGCTCTAGAAGACCTTTATCGCAATG 60
DB 2137324 GAAGTCGTTGCTCGCAAGGAAAGGAACCTGATGAGGCTCTAGAAGACCTTTATCGCAATG 2137265
QY 61 GCTTAAGAAATACGTCGTTTCAACCGTCGATTCGGCGGGGAAACGACGCTTTCTTTTGTCT 120
DB 2137264 GCTTAAGAAATACGTCGTTTCAACCGTCGATTCGGCGGGGAAACGACGCTTTCTTTTGTCT 2137205
QY 121 GCAAGAGTCTTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTCACATTTGTTA 180
DB 2137204 GCAAGAGTCTTTGGAAGATTTTTCGAAAAATGCTGGCAACATCAACAGTCACATTTGTTA 2137145
QY 181 GAAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGCC 240
DB 2137144 GAAACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGCC 2137085
QY 241 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 300
DB 2137084 CAAACCCAAAAATATGCGGGTCGAGATCTCAAAAGCTGCCATTTGCTGGGGATCGGACT 2137025
QY 301 GGGGTCTCTGTTCTTTGGGGAATGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 360
DB 2137024 GGGGTCTCTGTTCTTTGGGGAATGCTTAAGCCCATGGGGTGGTACATCTCTGTTGC 2136965
QY 361 AGTTTATGCTGCTGAGCAACATGGAAGTGGTAGCAGACTTAAAGAGCGGCTATCA 420
DB 2136964 AGTTTATGCTGCTGAGCAACATGGAAGTGGTAGCAGACTTAAAGAGCGGCTATCA 2136905


```
QY 421 TTTGCCACTGCCGATTATGATCATCGCGGTGAGGCAATCATCTGGTGTGTCAGGCGATT 480
Db 2136904 TTTGCCACTGCCGATTATGATCATCGCGGTGAGGCAATCATCTGGTGTGTCAGGCGATT 2136945
QY 481 TGGCAGATGGGCAATTTGGCGCTTTTGTGGGCACTGTTGGTGTGATGATTTCCG 540
Db 2136944 TGGCAGATGGGCAATTTGGCGCTTTTGTGGGCACTGTTGGTGTGATGATTTCCG 2136985
QY 541 AATTTCTACAAATGGCAGGAAAGAAAGCCGCAACTATTTGAGGACACCTCTGGG 600
Db 2136784 AATTTCTACAAATGGCAGGAAAGAAAGCCGCAACTATTTGAGGACACCTCTGGG 2136725
QY 601 CATCTTGGTGTGTCACCTCGGATTCATTTGTCGGAAGCTTCGCTCGGATGCTGCTGAT 660
Db 2136724 CATCTTGGTGTGTCACCTCGGATTCATTTGTCGGAAGCTTCGCTCGGATGCTGCTGAT 2136665
QY 661 GCAAAACAAATCCATCCGCGGTACATATTTGACGTTTATGCTGTGTGATCGC 720
Db 2136664 GCAAAACAAATCCATCCGCGGTACATATTTGACGTTTATGCTGTGTGATCGC 2136605
QY 721 ATCGGATGTGGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGGCGCGTT 780
Db 2136604 ATCGGATGTGGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGGCGCGTT 2136545
QY 781 GGTGAGTCCGAAAGTCTTGGGAAGCTTTGGCGGTCCATTTGCTTAGGATCGGTAC 840
Db 2136544 GGTGAGTCCGAAAGTCTTGGGAAGCTTTGGCGGTCCATTTGCTTAGGATCGGTAC 2136485
QY 841 TGGTGCACTAGTGTTCATCTTCGTCGATCAGTACACACTGGTGGATGCTGTGATTTGG 900
Db 2136484 TGGTGCACTAGTGTTCATCTTCGTCGATCAGTACACACTGGTGGATGCTGTGATTTGG 2136425
QY 901 TTGTGCCCTAGTGTGTGGCGCACGTTGGGTGACTTGGTGAAGTGAAGTCAAAACGGA 960
Db 2136424 TTGTGCCCTAGTGTGTGGCGCACGTTGGGTGACTTGGTGAAGTGAAGTCAAAACGGA 2136365
QY 961 TTTGGGATCAAGATATGTCGAACCTTCTTCCAGGCAACCGCGGATGATGACCGGTT 1020
Db 2136364 TTTGGGATCAAGATATGTCGAACCTTCTTCCAGGCAACCGCGGATGATGACCGGTT 2136305
QY 1021 GGATGGATGCTCCGCGCGGATGAGTGTGATCTGATGCTGATGATGATGATGATGATG 1080
Db 2136304 GGATGGATGCTCCGCGCGGATGAGTGTGATCTGATGCTGATGATGATGATGATGATG 2136245
QY 1081 GTATCCGFTGAAAGCTTGGCGGAGTTAAAGTTCAAAAGTTTAAAGGCTTTAAGGATTT 1140
Db 2136244 GTATCCGFTGAAAGCTTGGCGGAGTTAAAGTTCAAAAGTTTAAAGGCTTTAAGGATTT 2136185
QY 1141 CATACGTTGCACTGACGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTAAGGATTT 1200
Db 2136184 CATACGTTGCACTGACGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTAAGGATTT 2136125
QY 1201 TTTTAACTGGCGGCTGACTCAAACTGCGCAGCCCAACCAAGCGCCCAATATCAATG 1260
Db 2136124 TTTTAACTGGCGGCTGACTCAAACTGCGCAGCCCAACCAAGCGCCCAATATCAATG 2136065
QY 1261 CACCGGTAATGGCTGCTAGTAGGAACCGATTCGGCTGG 1300
Db 2136064 CACCGGTAATGGCTGCTAGTAGGAACCGATTCGGCTGG 2136025
```

RESULT 3

US-10-627-476-99
; Sequence 99, Application US/10627476
; Publication No. US20040030116A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schoder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE

; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125PCPN
; CURRENT APPLICATION NUMBER: US/10/627,476
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 09/602,787
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 99
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(955)
; OTHER INFORMATION: RXA01894
US-10-627-476-99

Query Match 75.2%; Score 978; DB 13; Length 978;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 136 AGAATTTTTCGAAAATGTGGCACCATCAACAGTGACATTTGTAGAACTTCAAGGAGA 195
Db 1 AGAATTTTTCGAAAATGTGGCACCATCAACAGTGACATTTGTAGAACTTCAAGGAGA 60
QY 196 ACCATGATGACCGGAGCAATCACCAGTCCATGAGGATGCCCAACCCAAATAA 255
Db 61 ACCATGATGACCGGAGCAATCACCAGTCCATGAGGATGCCCAACCCAAATAA 120
QY 256 TGGCGGTGAGATCTCAAAGCTGCCATTGCTGTGGGATCGGACTGGGGTCTTGGTTCT 315
Db 121 TGGCGGTGAGATCTCAAAGCTGCCATTGCTGTGGGATCGGACTGGGGTCTTGGTTCT 180
QY 316 TTTGGGATGCTTAAAGCCATGGGGTTGGTACATCTCTGTGAGGTTTATGGCTGC 375
Db 181 TTTGGGATGCTTAAAGCCATGGGGTTGGTACATCTCTGTGAGGTTTATGGCTGC 240
QY 376 AGCAACATGGAGTGTGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 435
Db 241 AGCAACATGGAGTGTGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 300
QY 436 TATGATCATCGCGGTGAGCAATCATCTGGTGTGATCGCCATTTGGCAGCATGGGCAT 495
Db 301 TATGATCATCGCGGTGAGCAATCATCTGGTGTGATCGCCATTTGGCAGCATGGGCAT 360
QY 496 TTTGGGCTCTTTTGTGGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
Db 361 TTTGGGCTCTTTTGTGGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 556 CACGAAAAGAACCCCGCAACTATTGAGGGACACCTCTGTGGGCACTTGTGTGTGTGTGT 615
Db 421 CACGAAAAGAACCCCGCAACTATTGAGGGACACCTCTGTGGGCACTTGTGTGTGTGTGT 480
QY 616 CTGGAATTCATTTGTTGGAGCTTCGCTGCGATTCGCTGCGATGCAAAACAATTCAT 675
```

Db 481 CTGGATTCATTTGTCGGAAGCTTTCCTGCGATGCTGCTGATGCAAAACATTCAT 540
 QY 676 CCGGGTACATATTCATTTGACGTTTCATGCTGTGTGTGATCGCATCGGATGCGGG 735
 Db 541 CCGGGTACATATTCATTTGACGTTTCATGCTGTGTGTGATCGCATCGGATGCGGG 600
 QY 736 GTATATCGCGGGTGTGCTTTTGGATCGACCCCAATGCGCGCTGTGTGAGTCCGAGAA 795
 Db 601 GTATATCGCGGGTGTGCTTTTGGATCGACCCCAATGCGCGCTGTGTGAGTCCGAGAA 660
 QY 796 GTCTTTGGGAAGCTTTTCCCGGCTCCATTTGTTAGGATCGGTCCTGCTGCACTCAGTGT 855
 Db 661 GTCTTTGGGAAGCTTTTCCCGGCTCCATTTGTTAGGATCGGTCCTGCTGCACTCAGTGT 720
 QY 856 TCACCTTCCTGCTCGATCACCACCTGCTGATGGGTGTGATCTTGGTTGTGCCCTAGTTGT 915
 Db 721 TCACCTTCCTGCTCGATCACCACCTGCTGATGGGTGTGATCTTGGTTGTGCCCTAGTTGT 780
 QY 916 GTGCGCCACGTTTGGGTGACTTGGTTGAGTTCGAGTTCAAACCGCAATTTGGGCATCAAGGA 975
 Db 781 GTGCGCCACGTTTGGGTGACTTGGTTGAGTTCGAGTTCAAACCGCAATTTGGGCATCAAGGA 840
 QY 976 TATGTCGAACCTTCTTCCAGCCACGCGGAGTGTGATGACCGTTTGGATGCGATGCTCCC 1035
 Db 841 TATGTCGAACCTTCTTCCAGCCACGCGGAGTGTGATGACCGTTTGGATGCGATGCTCCC 900
 QY 1036 GCGCCGATGTTGAGTGTGATCTGATCTGAGTGTGATCAGCAGCTGCTATCCGTCGTAAG 1095
 Db 901 GCGCCGATGTTGAGTGTGATCTGATCTGAGTGTGATCAGCAGCTGCTATCCGTCGTAAG 960
 QY 1096 CTGCGCCACGTTTAAAGT 1113
 Db 961 CTGCGCCACGTTTAAAGT 978

RESULT 4

US-09-738-626-2216
 ; Sequence 2216, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENO, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 2216
 ; LENGTH: 855
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-2216

Query Match 65.8%; Score 855; DB 9; Length 855;
 Best Local Similarity 100.0%; Pred. No. 7.9e-275; Indels 0; Gaps 0;
 Matches 855; Conservative 0; Mismatches 0;

QY 236 ATGCCCAACCCAAAATAATGCGGTCGAGATCTCAAAGCTGCCATTGCTGTGGGATC 295
 Db 1 ATGCCCAACCCAAAATAATGCGGTCGAGATCTCAAAGCTGCCATTGCTGTGGGATC 60
 QY 296 GGAAGTGGGGTCTGCTGTTCTTTTGGGATTTGCTAAAGCCCATGGGTTGGTACATCCTC 355
 Db 61 GGAAGTGGGGTCTGCTGTTCTTTTGGGATTTGCTAAAGCCCATGGGTTGGTACATCCTC 120
 QY 356 GTTGCAAGTTTATGCTGTCAGCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGGCG 415
 Db 121 GTTGCAAGTTTATGCTGTCAGCAACATGGAAGTTGGTAGCAGACTTAAAGAGCGGCG 180
 QY 416 TATCATTTGCCACTGCGGATTTATGATCATCGGCGGTCAGCAATCATCTGGCTGTCTATG 475
 Db 181 TATCATTTGCCACTGCGGATTTATGATCATCGGCGGTCAGCAATCATCTGGCTGTCTATG 240
 QY 476 CCATTTGGCAAGTGGGCAATTTTGGGCTCTTTTGTGGCCACTGTGTGTGTGTGTAT 535
 Db 241 CCATTTGGCAAGTGGGCAATTTTGGGCTCTTTTGTGGCCACTGTGTGTGTGTGTAT 300
 QY 536 TTCGGAATTTCTACATCGCAGGAAGAGCCGCAACTATTTGAGGACACCTCT 595
 Db 301 TTCGGAATTTCTACATCGCAGGAAGAGCCGCAACTATTTGAGGACACCTCT 360
 QY 596 GTGGGCAATCTTGTGCTCACCTGGATTCATTTGCGAAGCTTCGCTCGGATGCTGTG 655
 Db 361 GTGGGCAATCTTGTGCTCACCTGGATTCATTTGCGAAGCTTCGCTCGGATGCTGTG 420
 QY 656 CTGATCAAAACAAATTCATCCCGGGTACATATTTCAATTTGACGTTTATGCTGTGTG 715
 Db 421 CTGATCAAAACAAATTCATCCCGGGTACATATTTCAATTTGACGTTTATGCTGTGTG 480
 QY 716 ATCGCATCGGATGCGGCTATATCGGCGGTATATCGGCGGTGTGTTTGGATCGCACCATGGG 775
 Db 481 ATCGCATCGGATGCGGCTATATCGGCGGTATATCGGCGGTGTGTTTGGATCGCACCATGGG 540
 QY 776 CCGTTGGTGAAGTCCGGAAGTCTTGGGAAGCTTTGCGGCTCCATTTGCTTAGGATCG 835
 Db 541 CCGTTGGTGAAGTCCGGAAGTCTTGGGAAGCTTTGCGGCTCCATTTGCTTAGGATCG 600
 QY 836 GTCACTGTGCACTGAGTGTTCACCTTCCTGCTCGATCACCACCTGCTGATGCTGTGATC 895
 Db 601 GTCACTGTGCACTGAGTGTTCACCTTCCTGCTCGATCACCACCTGCTGATGCTGTGATC 660
 QY 896 TTGGGTTGCCCCCTAGTTGTGTGCGCCAGCTTTGGGTGATCTTGGTTGAGTTCAGTTCAA 955
 Db 661 TTGGGTTGCCCCCTAGTTGTGTGCGCCAGCTTTGGGTGATCTTGGTTGAGTTCAGTTCAA 720
 QY 956 CCGGATTTGGGCATCAAGGATATGCGAACCTTTCTCCAGGCCACGCGGATGATGAGC 1015
 Db 721 CCGGATTTGGGCATCAAGGATATGCGAACCTTTCTCCAGGCCACGCGGATGATGAGC 780
 QY 1016 CGTTTGGATGGATGCTCCCGCCCGGATGATGAGTGTGATCTTGGTTGATCTGATGATCAGC 1075
 Db 781 CGTTTGGATGGATGCTCCCGCCCGGATGATGAGTGTGATCTTGGTTGATCTGATGATCAGC 840
 QY 1076 AGCTGTATCCGTCG 1090
 Db 841 AGCTGTATCCGTCG 855

RESULT 5

US-10-282-122A-17786
 ; Sequence 17786, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Walli, Daniel

```
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 17786
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-17786

Query Match      26.3%; Score 342.2; DB 13; Length 876;
Best Local Similarity 63.1%; Pred. No. 4e-103;
Matches 527; Conservative 0; Mismatches 308; Indels 0; Gaps 0;

QY 237 TGCCCAACCCAAATAATCGGGTCAGATCTCAAGCTGCCATGCTCTGGGGATCG 296
DB 32 TGCCCTAAACCGAAGAAATTCGGCTGGTCGAACCTGAAAGCCGCGATCAGCCTTGGAAATG 91

QY 297 GACTGGGGGCTCTGCTTTCTTTGGGGATTGCTCTAAAGCCCATGGGGTTGGTACATCCTCG 356
DB 92 GCTTGGGCGCCCTCGATTGCTCGTATCTTTGTTATTCCTTTTGGTTGGTATCCATTAG 151

QY 357 TTGCAGGTTTATGGCTCAGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGCGCT 416
DB 152 TGGCTATCGCAATCGCAGTGGCAACTTGGGAGGTAGAGCGCGTTTAAATAGAAGCCGGTT 211

QY 417 ATCATTTCCCACTGCGATTATGATCATCGCGGTGAGGCAATCATCTGGCTGTCATGGC 476
DB 212 ATCTTTTCAACGTTGGGTGATGCTTATGGGACAAAGTCACTGTTGGCTTAGTTGGC 271

QY 477 CATTTGGCAGCATGGCAATTTTGGCGCTTTTGGGCCACTGTGTGGTGTCTGATGATT 536
DB 272 CATTTGGTCCGAAGGATTAGTGGCTGTTTCGTCGGGGTTGTACTCGCTACGATGTTG 331

QY 537 TCCGAATTTTCTACATGGCAGGNAAGAACCCGCAACTATTAGGACACCTCTG 596
DB 332 GTCGCGCTTTTCCATCATGGTCTGTTATGSCCCAGAAAACTATCTCGGAGATACTGCGG 391

QY 597 TGGGCATCTCTGCTCACCCTGGATTCCATGTTTCGGAAGCTTCGCTCGCATGCTGCGC 656
DB 392 TCGCAATTTTGTTTTAACTGGATCCCACTTTTGGTAGTTTTCGSCATGCTTTCCG 451

QY 657 TGATCCAAAACAATTCATCCCGGGTACATATTTCAATTTGACGTTATGCTGTGTGTA 716
DB 452 TGTTTGAACACAGAAACAGACCCGGAAGTAGTCTTCATGTCACATTCATGTTGTGCGTTA 511
```

RESULT 6

US-10-282-122A-26399
Sequence 26399, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: Patentin version 3.1

SEQ ID NO 26399

LENGTH: 918

TYPE: DNA

```

; ORGANISM: Mycobacterium bovis
US-10-282-122A-26399

Query Match      14.8%; Score 191.8; DB 13; Length 918;
Best Local Similarity 54.4%; Pred. No. 9.3e-53;
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;

QY 257 GCGGTCGAGATCTCAAAGCTCCCATCTGCTGGGGATCGGACTGGGGTCTCTGTTCTT 316
DB 88 GCGGCGCGTATTGCGGCACGATCTGTTGGGCTTCCATAGGCTAGTCTCTATC 147
QY 317 TTGGGATGTCCTAAGCCCATGCGGCTTGGTACATCTCGTTGCAAGTTTATGGCTGCA 376
DB 148 GCGGTCGCTGCTCGTTCGCGCGCTTGGGTGGCATGCTGCGCGCTCGCCACCTTGGTC 207
QY 377 GCAACATGGGAATGGTAGCAGACTTAAAGAGCGGCTATCATTTGCGCATGCGGAT 436
DB 208 GCTACCCATGAGTGTGCGGAGTTGCGGAGGCGGCTATCTCATCCGGTTATCCCG 267
QY 437 ATGATCATCGCGGTCAGGCAATCATCTGCGCTGTCATGCGCATTTGGCACGATGGCAT 496
DB 268 TTGCTGATTTGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
QY 497 TTGGGCTCTTTTGTGGCCACTGTTGGTGTGATATTTCCGAATTTTACAAATGGC 556
DB 328 TTGGCGGGCTTTGGTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
QY 557 ACGGAAAAAGAGCCC-----GCACTATTGAGGGACACC 592
DB 388 AGCGTGACGCGCGCGGACGACCGCGGTGACCGTCCCGGGAACTACTTGTCCGACGTC 447
QY 593 TCTGTGGGCACTTCTGCTGCTACCTGGATTCATTTGCGAAGCTTCTGCTGCTGCTGCTG 652
DB 448 TCGGCCACGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507
QY 653 TCGCTGATGCAAAACAATTCATCCCGGTATATTTTCACTTTGACGTTTCACTGCTGCT 712
DB 508 GTCATCCCGGAAATGGCTCGGATGG-----TGTTCTGATGATGATCGCG 555
QY 713 GTGATCGCATCGGATGTTGGGGGATATATCGCGGTGTTTCTTTGGATCGCACCAATG 772
DB 556 GTCATCGCTTCGATGTCGCGGCTACGCGGTGGGGTGTCTTTGGCAAGCATCCGATG 615
QY 773 GCGCGTGTGTCAGTCCGAGAGCTTTGGAGAGCTTTGCGGCTCCATGCTTATGGA 832
DB 616 GTTCCGACGATCAGCCCGAAGAGTCTGTGGAGGCTTTGCGGGTCTGCTGCTGCTGCGG 675
QY 833 TCGGTCACCTGGTCACTCAGTCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
DB 676 ATCACCAGCAACGATCATCACCGGACTTTCTTGTGCGCAAAACGCGTGGATTGGTGA 735
QY 893 ATCTTGGTGTGCGCTAGTGTGTGCGGACAGTTGGGTGCTGCTGCTGCTGCTGCTGCTG 952
DB 736 CTGCTCGCGCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
QY 953 AAACCGGATTTGGGATCAAGATATGTCGAACCTTTCTCAGGCGCACGCGGATTTGATG 1012
DB 796 AAACGTGACCTCGGATCAAGACATGAGCGGCTGCTTACCGGCCACGCGGCTGCTGAT 855
QY 1013 GACCGTTTGGATGTCATGCTCCCGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1071
DB 856 GACCGGCTCGACGGATGCTGCTTCCGCGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 914

RESULT 7
US-10-282-122A-28605
; Sequence 28605, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Hasei,beck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELI/PA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28605
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28605

Query Match      14.8%; Score 191.8; DB 13; Length 921;
Best Local Similarity 54.4%; Pred. No. 9.3e-53;
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;

QY 257 GCGGTCGAGATCTCAAAGCTCCCATCTGCTGGGGATCGGACTGGGGTCTCTGTTCTT 316
DB 88 GCGGCGCGTATTGCGGCACGATCTGTTGGGCTTCCATAGGCTAGTCTCTATC 147
QY 317 TTGGGATGTCCTAAGCCCATGCGGCTTGGTACATCTCGTTGCAAGTTTATGGCTGCA 376
DB 148 GCGGTCGCTGCTCGTTCGCGCGCTTGGGTGGCATGCTGCGCGCTCGCCACCTTGGTC 207
QY 377 GCAACATGGGAATGGTAGCAGACTTAAAGAGCGGCTATCATTTGCGCATGCGGAT 436
DB 208 GCTACCCATGAGTGTGCGGAGTTGCGGAGGCGGCTATCTCATCCGGTTATCCCG 267
QY 437 ATGATCATCGCGGTCAGGCAATCATCTGCGCTGTCATGCGCATTTGGCACGATGGCAT 496
DB 268 TTGCTGATTTGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
QY 497 TTGGGCTCTTTTGTGGCCACTGTTGGTGTGATATTTCCGAATTTTACAAATGGC 556
DB 328 TTGGCGGGCTTTGGTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
QY 557 ACGGAAAAAGAGCCC-----GCACTATTGAGGGACACC 592
DB 388 AGCGTGACGCGCGCGGACGACCGCGGTGACCGTCCCGGGAACTACTTGTCCGACGTC 447
QY 593 TCTGTGGGCACTTCTGCTGCTACCTGGATTCATTTGCGAAGCTTCTGCTGCTGCTGCTG 652
DB 448 TCGGCCACGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507
QY 653 TCGCTGATGCAAAACAATTCATCCCGGTATATTTTCACTTTGACGTTTCACTGCTGCT 712
DB 508 GTCATCCCGGAAATGGCTCGGATGG-----TGTTCTGATGATGATCGCG 555
QY 713 GTGATCGCATCGGATGTTGGGGGATATATCGCGGTGTTTCTTTGGATCGCACCAATG 772
DB 556 GTCATCGCTTCGATGTCGCGGCTACGCGGTGGGGTGTCTTTGGCAAGCATCCGATG 615
QY 773 GCGCGTGTGTCAGTCCGAGAGCTTTGGAGAGCTTTGCGGCTCCATGCTTATGGA 832
DB 616 GTTCCGACGATCAGCCCGAAGAGTCTGTGGAGGCTTTGCGGGTCTGCTGCTGCTGCGG 675
QY 833 TCGGTCACCTGGTCACTCAGTCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
DB 676 ATCACCAGCAACGATCATCACCGGACTTTCTTGTGCGCAAAACGCGTGGATTGGTGA 735
QY 893 ATCTTGGTGTGCGCTAGTGTGTGCGGACAGTTGGGTGCTGCTGCTGCTGCTGCTGCTG 952
DB 736 CTGCTCGCGCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
QY 953 AAACCGGATTTGGGATCAAGATATGTCGAACCTTTCTCAGGCGCACGCGGATTTGATG 1012
DB 796 AAACGTGACCTCGGATCAAGACATGAGCGGCTGCTTACCGGCCACGCGGCTGCTGAT 855
QY 1013 GACCGTTTGGATGTCATGCTCCCGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1071
DB 856 GACCGGCTCGACGGATGCTGCTTCCGCGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 914

```

```
QY 653 TCGCTGATGCAAAACATTCATCCCGGTACATATTTCAATTTGAGCTTCATGCTGTGT 712
Db 508 GTCTACCCGGAAATGCTCGGATGG-----TGTCTGATGATGATCGCG 555
QY 713 GTGATCGCATCGGATGCGGGGTATATCGGGGTGTGTTCTTTGGATCGCACCAATG 772
Db 556 GTCATCGCTTCGATGTCGCGGCTACCGCTGGGGGTGCTGTTTGGCAAGCATCCGATG 615
QY 773 GCGCGGTGTGTGAGTCGGAAGACTCTGGAAAGCTTTTCCCGCTCCATTTGCTTAGGA 832
Db 616 GTTCCGACGATCACCAGAAAGTCTGGAGGGCTTTTCCCGTTCGCTGTGTGCGGG 675
QY 833 TCGGTCACTGTGCACTAGTTCATCTTCTGCTCGATCACCACCTGGTGTGATGGTGTG 892
Db 676 ATCACCAGCAACGATCATCACCGCACTTCTCTGTTCGGCAAAAGCGCGTGGAT 735
QY 893 ATCTTGGGTGTGCTTGTGTCGCGCACGTTGGTGTGACTTGTGTGATGCGAGTTC 952
Db 736 CTGCTCGCGGTGCTTTTCTGCTCACCACCGCTGGGACCTGTGTGGATGCGAGTTC 795
QY 953 AAACGCGATTTGGGCATCAAGGATATGTCGAACCTTCTTCCAGCCACGCGGATGATG 1012
Db 796 AAACGTGACCTCGCATCAAGACATGGGCGCGCTGCTACCGGCCACGCGGTCTGATG 855
QY 1013 GACCGTTGGATGATGCTCCCGCGCGATGCTGAGCTGTTGATCCTGAGTGTGAT 1071
Db 856 GACCGGCTCGACGCATACCTGCTTCCCGGTGGCGCTGGATGATGCTCCTCACACTGCT 914
```

```
RESULT 8
US-10-282-122A-25992
; Sequence 25992, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIORITY FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 25992
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Mycobacterium avium
US-10-282-122A-25992

Query Match      13.0%; Score 158.8; DB 13; Length 933;
Best Local Similarity 52.8%; Pred. No. 4.7e-45;
Matches 439; Conservative 0; Mismatches 357; Indels 36; Gaps 2;

QY 257 GCGGTCGAGATCTCAAAGCTGCCATTTGCTGTGGGGATCGGACTGGGGTCTCTGTTCTT 316
Db 103 GCGGACGTGACCTCGGGGCCCATCGCGTGGGCGCGGTATCGGGCCGCTGCTCATC 162
QY 317 TTGGGATTTGCTTAAAGCCATGGGGTGTGTACATCTCTGTTGACAGTTTATGCTGCA 376
Db 163 GTACGCTGTGTTTCCGGCCCGCTTCTGGTGGCGATCGTGGGATGCCCATCTCTGTTC 222
QY 377 GCAACATGGGAAGTTGGTAGCAGACTTAAAGAGGGCGCTATCATTTGCCACTGCCGATT 436
Db 223 GCAGCCATGAGTGTTCGCGCGCTGCGGAGGCGGATACGTATCCGGTATTCGG 282
QY 437 ATGATCATCGGCGGTACGCAATCATCTGCTGTGTATGSCCATTTGGACAGATGGCAT 496
Db 283 TTGCTGGCGCGGCGGACGCTCAAGTGTGCTGACCTGGCCGTTCCACGCCCGCGCGG 342
QY 497 TTGGCTCTTTTGTGGCCACTGTTGGTGTGTATGATGATTTCCGAATTTTCTACATGTC 556
Db 343 TTGGCGGCTTCCGGCTCACCGTGTGGCTGCTGTTCTTGGCGGTGTTTCATGAGGAC 402
QY 557 ACGGAAAAAGAGCCCGC-----AACTATTGAGGGACACC 592
Db 403 AACCGCAAGCGCCCGAGCGCTTGGCGGTTCGCCCTCGCGCACTACCTGGCGACGGG 462
QY 593 TCTGTGGCATCTTCTGTCTCACTGGATTCATTTGTTGGAGCTTCTGCGGATGTCGTCG 652
Db 463 TCGGCCACCGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522
QY 653 TCGCTGATGCAAAACAAATTCATCCCGGTACATATTTTCATTTTGACGTTGATGCTGT 712
Db 523 -----GTCTATCCGCGCGAGGTGCGGGCGGGTGTTCGCTGATGATCACC 570
QY 713 GTGATCGCATCGGATGCGGGGTATATCGCGGGTGTGTTCTTTGGATGCGACCAATG 772
Db 571 GTGGTGGCTCCGACGCTCGCGGCTACGCGGTGCGGTGCTGTTCGCGCAAGCATCCGATG 630
QY 773 GCGCGCTTGTGTGAGTCCGAGAGTCTTGGAGAGCTTTGCGGGCTCCATTTGTTTAGGA 832
Db 631 GTCCCGCGGATCAGCCCCAAGAAATCTGGAGGGCTCGCGGCTCGCTGTGCTGGGT 690
QY 833 TCGGTCACTGCTCACTCACTGTTTCACTTCTCTCTCGATCACACATGTTGGATGGTGTG 892
Db 691 ATCACCAGCGCCACCTCGCGCGGACTTTTCTGGCGCAAGCGCGCTGGGTGGCGGG 750
QY 893 ATCTTGGTGTGCTTGTGTTGTGTCGCGCATGTTGGTGTGACTTGTGTTGAGTGCAGTTC 952
Db 751 CTGCTGGGGTGTGTTGTTGTTTCACTGACCTCGCGGACCTGTGGAGTGCAGAGTC 810
QY 953 AAACGCGATTTGGGCATCAAGGATATGTCGAACCTTCTTCCAGGCCACGCGGATGATG 1012
Db 811 AAGGTGACCTGGGCATCAAGGATGAGGCGGCTGCTGCGCGCACGCGCGCTGATG 870
QY 1013 GACCGTTGATGCGATGCTTCCCGCGCGGATGTTGACGTTGTTGATCCTGA 1064
Db 871 GACCGGCTGAGACGCTGCTGCTGCGGTGCGCGCTGCGCGCTGGAGCGGTGCTGA 922
```

```
RESULT 9
US-10-282-122A-27700
; Sequence 27700, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA 034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 27700
 ; LENGTH: 939
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium leprae
 ; US-10-282-122A-27700

 Query Match 12.6%; Score 164; DB 13; Length 939;
 Best Local Similarity 52.6%; Pred. No. 1.9e-43;
 Matches 440; Conservative 0; Mismatches 375; Indels 21; Gaps 3;

 QY 257 GCGGTCGAGATCTCAAGCTGCCATTCCTGTGGGATCGGACTGGGGGCTCTGTTCCTT 316
 DB 97 GCGGACGTAATCTGCCCGCGCGATCGCGTGGGCTTAAGTATGGTGTCTCGTC 156

 QY 317 TTGGGATGTCTTAAGCCCATGGGTGTGATCTCTGTGACAGTCTTTATGGCTGCA 376
 DB 157 GCAACGTGGTGTTCCTCCGGAATCTGGTGTCTGTGCGCTGGCCATTTTCCTT 216

 QY 377 GCAACATGGGAAGTGTGAGAGACTTAAGAACGGCGCTATCATTTGCCACATCCGGA 436
 DB 217 GCTAGCCATGAGGTGGTGGCGCGCTACGGGAAGCTGGATATGATTCGGGCTATCCCG 276

 QY 437 ATGATCATCGCGGTGAGGCAATCATCTGCTGTCATGGCCATTTGGCAGATGGGCATT 496
 DB 277 CTGCTCATCGTGGGAGTTTACCGTCTGGTTAACTTGGCCGCTATCGCACCGTGTGGA 336

 QY 497 TTGCGGTCTTTTGTGGCAGTGTGTGGTGTGATGATTTCCGAATTTT-----CTAC 550
 DB 337 TTGCTGTGCTTCGGTGGCAGGTGGTGTCTGCAATGATTTGGCGCTTGTCAAGCAGAC 396

 QY 551 AATGGCAGGAAAGAACCGCAACTATTAGGGACACCTCTGTGGGCATCTTCGIG 610
 DB 397 AACAGCAGACGACGAATCTCGGGAAGCTTTGGCAGGTCCCGGCTGCAAAATTACCTG 456

 QY 611 -----CTCACCTGATTCCTATTGTTGCGAAGCTTCGCTGGATGCTGCTCATGACAA 665

DB 457 CCGGACGCGTCGGCCACCGTTTTTTTGGCGCGGTGGTCCCGTTGTTGCTCGTTCGCA 516
 QY 666 ACA-----ATTCCATCCCGGTACATATTTTCAATTTTACGTTTATGCTGTGTGTG 715
 DB 517 GCATTGCTGTTTATCCGAAGATGTGCGGGCGGGTATTCTGCTGATGATTCGGGTG 576
 QY 716 ATCGCATCGGATGTGGCGGGTATATCGCGGGTGTGTTCTTTGGATCGCACCAATGGCG 775
 DB 577 GTAGCTCCGATGTGGGTGTTACACAGTAGGGTACTTTTCGGCAACATCCACTGGTT 636
 QY 776 CCGTTGGTGGTCCGAAGTCTTGGGAAGCTTTGCCGGCTCCATTTGCTTAGGATCG 835
 DB 637 CCAAGGATTAGCCCGAATAAGTCTCGGAGGATTCGCCGGTTCGCTGTCTCGCGGACC 696
 QY 836 GTCACCTGGTGCACCTCAGTGTTCACCTTCCTGCTCGATCACCACTGGTGGATGGTGTGATC 895
 DB 697 ACCGAACCATCTGACTGCTACCTTCCTGGCTGGCAACACCGTGGGTTCGGGCTTGG 756
 QY 896 TTGGGTGTCCCTAGTGTGTGCGGCACGTGGGTGACTTGGTTGAGTCGCAATTCAAA 955
 DB 757 CTCAGTTTTGTACTGTGTCTCACCTGCACGCTGGGGATTTGGTGGAGTCCCAAGTTAAG 816
 QY 956 CGCATTTGGGCATCAAGGATATGCGAACCTTCTTCCAGGCCACGGCGGATTTGATGGAC 1015
 DB 817 CGGACCTCGGCATCAAGATATGGCCGGCTGTACCCGGCCACGGTGGCTGTATGAC 876
 QY 1016 CGTTTGGATGGCATGCTCCGCGCGGATGGTGAAGTGTGATCTGAGTGTGAT 1071
 DB 877 CGCTTGTATGTTGCTGCTCGCGTGGCGGTGTAGCTTGGACCATGCTTACGTTGCT 932

 RESULT 10
 US-09-864-408A-905
 ; Sequence 905, Application US/09864408A
 ; Publication No. US2004000947A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leach, Martin D.
 ; APPLICANT: Shinkets, Richard A.
 ; TITLE OF INVENTION: No. US2004000947A1 Human Polynucleotides and Polypeptides Encod
 ; FILE REFERENCE: 21402-012
 ; CURRENT APPLICATION NUMBER: US/09/864,408A
 ; CURRENT FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: 60/206,690
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 9068
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 905
 ; LENGTH: 297
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-864-408A-905

 Query Match 9.5%; Score 123.4; DB 11; Length 297;
 Best Local Similarity 68.0%; Pred. No. 3.7e-30;
 Matches 172; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

 QY 854 GTTCACTTCTCTCGATCACCATGTGGTGGATGGTGTGATCTTGGGTGTGCCCTAGTT 913
 DB 1 GTGCACTTCTCATCCAGGACCTTGTGGATGGCGTGTGCTGGGGGTTCGCGTTGTC 60

 QY 914 GTGTGGCCACGTGGGTGACTGTGTTGAGTCGCGATTTCAACCGGATTTGGGATCAAG 973
 DB 61 CTCTCGCCACCATGGCGACCTCTGTGAGAGCCAGTTCAAGCGCGAGCTGGGAATCAAG 120

 QY 974 GATATGTGGAACCTTCTTCCAGGCCACCGCGGATTTGATGGACCGTTTGGATGGCATGCTC 1033
 DB 121 GACATGTGAACTTCTTCCGGCCACCGCGGGCTGTAGACCCCTTGGACCGGATGCTG 180

 QY 1034 CCGGCGCGATGGTGAAGTGTGATCTGATCTGATGATGATGATGATGATGATGATGATGAT 1093
 DB 181 CCGCGCGCAGCGGACCTTACATGCTGCTCAACCGCGCGGATCATGAACCCCTGTAG 240

QY 1094 AGCTTGGCCAGC 1106
Db 241 CGGCCCTGGCAGC 253

RESULT 11

US-09-738-626-2215/c
; Sequence 2215, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 2215
; LENGTH: 258
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-2215

Query Match 9.0%; Score 117; DB 9; Length 258;
Best Local Similarity 100.0%; Pred. No. 4.7e-28;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1184 AAGGCTTTACGGATTTTCTTAAGTGGCGAGGTACTCAACATGGCGCAGCCACCAAC 1243
Db 258 AAGGCTTTACGGATTTTCTTAAGTGGCGAGGTACTCAACATGGCGCAGCCACCAAC 199
QY 1244 AAGGCCATATCAATGACCGGTAATGGCTAGTAGGAAACCGATTCCGGCTGG 1300
Db 198 AAGGCCATATCAATGACCGGTAATGGCTAGTAGGAAACCGATTCCGGCTGG 142

RESULT 12

US-10-156-761-2610
; Sequence 2610, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2610

; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1176)
US-10-156-761-2610

Query Match 8.5%; Score 110.2; DB 15; Length 1176;
Best Local Similarity 54.6%; Pred. No. 2.3e-25;
Matches 242; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 624 CATTGTCGGAAGCTTCGGTGGATGCTGCTGCTGATGCAAAACAATTCATCCCGGTA 683
Db 713 CGTTCTAGTCCCGTTCTCGGACGTTCTGCTGCGATGATGCTACGGCCGACGAGTTC 772
QY 684 CATATTTCAATTTGACGTTTCATGCTGTGTGATCGCATCGGATGCGGCGGTATATCG 743
Db 773 CGCGCGGGTGTCTCAGTTCTCTGCTGCTCACCGTGTAGCGACACGGGCGGTACGCGA 832
QY 744 CGGGTGTGTTCTTTGGATCGCACCAATGGCCCGTTGGTGTAGTCCGAAGAAAGTCTTGGG 803
Db 833 TCGGCTGGCGTTTCGGCAAGCACAGGCTCGCCCGCGCATCAGCCCGGCAAGACCCGCG 892
QY 804 AAGGCTTTCGGCTCCATTGCTTAGGATCGTCACTGGTCACTCAGTGTCTCACTTCC 863
Db 893 AGGGCTGTGTCGGAGCGGTCTCGTTTCGGCATGGTGGCGGCGCGCTGTGCAATGGAGTTC 952
QY 864 TGCTCGATCACCA---CTGGTGGATGGGTGTGATCTTGGGTTGTGCTTGTGTGG 920
Db 953 TGATCGACGACGGCAGCTGTGGCAGGCGCTGTGCTCGGCTTCGGGTTCGCCCGCAGTG 1012
QY 921 CCAGCTTGGGTGACTTGGTGTGAGTTCGAGTTCAAGCGGATTTGGGCATCAAGGATATGT 980
Db 1013 CCAGCTTCGGGACCTTCGGCGAGTCCATGATCAAGCGGATCTGGGCATCAAGGACATGG 1072
QY 981 CGAACCTTCTTCAGGCCACGGGATTCATGACCGTTTGGATGCGATGCTCCCGGCG 1040
Db 1073 GCAGCTGTGTCGGGCGACGGCGCATGATGACCGGCTGACTCGTCTGCTGCTGCGCGCG 1132
QY 1041 CGATGGTACGCTGGTTGATCCTG 1063
Db 1133 CTCGGTGGTGGTTGTTGCTG 1155

RESULT 13

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)

OTHER INFORMATION: a, t, c, g, other or unknown

```

US-10-156-761-1
Query Match      8.5%; Score 110.2; DB 15; Length 9025608;
Best Local Similarity 54.6%; Pred. No. 5.5e-23;
Matches 242; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 624 CATTTGTCGGAAGCTTCGCTCGATGCTGTCTGCTCATGCAAAACAAATTCATCCCGGGA 683
DB 3223095 CGTTCTACGTCCTCGGTCCTGGGACAGTTCGTCGCGATGATCTCAGGCCACGACGGTC 3223036

QY 684 CATATTTCATTTGACGTTTCATGCTGTGTGTGTGATCGCATCGATGCGGCGGTATATCG 743
DB 3223035 CGCGCGGGTGTCTACGTTCTCTGCTGCTCACCGTGTGAGCGACACGCGCGCTACGGCA 3222976

QY 744 CGGGTGTGTCTTTTGGATGCGACCAATGCGCGCGCTTGGTGTGAGTCCGAGAGTCTTGGG 803
DB 3222975 TCGGCTGGGTTTCGCAAGCACAGGCTCGCCCGCGCATCAGCCCGCGCAAGACCGCGG 3222916

QY 804 AAGGCTTTCCGGCTCCATTGTCTTAGGATCGGTCACTGTGTCACATCAGTGTTCACATTCC 863
DB 3222915 AGGGCTGTGTCTCGAGCGGTCTCGTTTCGCGATGTTGGCGGCGCGCTGTGATGAGTTCC 3222856

QY 864 TGCTCGATCACACA---CTGTGGATGGTGTGATCTTGGTGTGCTTGGCTTGTGTGCG 920
DB 3222855 TGATCGACGACGGCAGCTGTGTGGCAGGCGCTCGTCTCGGCTTCGCGGTCCGCCCAAGT 3222796

QY 921 CCACGTTGGGTGACCTTGTGTGAGTCGCACTGTTCAACACGCGATTTGGGATCAAGGATATGT 980
DB 3222795 CCACGCTCGCGACCTCGCGAGTCCATGATCAGCGGATCTGGGCATCAAGGACATGG 3222736

QY 981 CGAACCTTCTTCAGGCGACGCGGATGATGAGACGCTTTGGATGGCATGCTCCCGCGCG 1040
DB 3222735 GCACGCTGTGCGCGGCGACGCGGCGCATGAGACGCGCTGCTGCTGCGCAGCG 3222676

QY 1041 CGATGCTGACGCTGTTGATCTGT 1063
DB 3222675 CTCGGTGTGTGTGTGTGCTG 3222653

RESULT 14
US-10-470-565-1
; Sequence 1, Application US/10470565
; Publication No. US20040126870A1
; GENERAL INFORMATION:
; APPLICANT: Societe des Produits Nestle S.A.
; TITLE OF INVENTION: NCC2705 - the genome of a Bifidobacterium
; FILE REFERENCE: 80290/WO
; CURRENT APPLICATION NUMBER: US/10/470,565
; PRIOR FILING DATE: 2003-07-29
; PRIOR APPLICATION NUMBER: EP 01102050.0
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 2256646
; TYPE: DNA
; ORGANISM: Bifidobacterium longum
US-10-470-565-1

Query Match      5.4%; Score 69.8; DB 17; Length 2256646;
Best Local Similarity 52.0%; Pred. No. 7.9e-10;
Matches 192; Conservative 0; Mismatches 162; Indels 15; Gaps 1;

QY 721 ATCGGATGTGGCGGTATATCGCGGTGTGTCTTTGGATCGCACCACCAATGCGCGGT 780
DB 1884103 ATCCGATACGGCGGTGTGTGTTCCGCTGCGCTCGGCAACACAAAGCTACGCCCGG 1884162

QY 781 GGTGAGTCCGAAGAGCTTTGGGAAGGCTTTGGCGGCTCCATTTGCTTAGGATCGGTCAAC 840
DB 1884163 CATCTCGCGAAGAGCTTTGGGAGGCTTTGGGCGCTTGTGCGCATGCTCTTCGCCATGCTG 1884222

QY 841 TGGTGCACTCAGTGTTCCTCTCTGCTGATCAC-----CACTGGTGGAT 895

```

RESULT 15

US-10-282-122A-33608
 ; Sequence 33608, Application US/10282122A
 ; Publication No. US20040029129A1

GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 33608

; LENGTH: 813

; TYPE: DNA

; ORGANISM: Pseudomonas syringae

US-10-282-122A-33608

Query Match 5.3%; Score 68.4; DB 13; Length 813;

Best Local Similarity 51.2%; Pred. No. 1.8e-11;

Matches 212; Conservative 0; Mismatches 196; Indels 6; Gaps 2;


```
Qy 640 CGCTGCGATGCTGCTGATGCAABACAATTCCATCCCGGTACATATTTTCATTTTGAC 699
Db |||
Qy 360 CGCGTGGCAGGGGTGGTGTCTCATCAGCAATGGCCGCTGGGCACTGGCTGATCCTGTC 419
Db |||
Qy 700 GTTCATGCTGTGTGTGATCGCATCGGATGTGGCGGGTATATCGCGGGTGTGTTCTTTGG 759
Db |||
Qy 420 GGTGATGGTGTGTGTGGCGGCTGACATCGGCGCGTATTTCTCGGCAAGGCTTTCGG 479
Db |||
Qy 760 ATCCACCCCAATGGCGCGTGTGGTGTGAGTCCGAGAAAGTCTTGGGAAGGCTTTCGCGCTC 819
Db |||
Qy 480 CAGCGCAAGCTGGCTCCCAAGSTCACTCCCGGTAAAGCTGGAGGGCGGTATACGGCGG 539
Db |||
Qy 820 CATTGTCTTAGGATCGG---TCACTGGTGCACTCAGTGTTCACCTTCCTGCTCGATCACCA 876
Db |||
Qy 540 TCTGGTGTGAGTCTGGGTATCACCGCGGCAGTCGGTGTGCGGTGACTGGACTGTGCT 599
Db |||
Qy 877 CTGGTGGATGGGTGTGATCTTGGGTGTGGCTTAGTGTGTGGG---CCAGTGTGGTGA 933
Db |||
Qy 600 TCAGTTCATCGCCGCACTGTTGGGTGCGGCTGTGATGCTGTTTCATTTCGCTGATCGGTGA 659
Db |||
Qy 934 CTTGGTTGAGTCGCAGTTCAAACCGGATTTGGGGCATCAAGGATATGTCGAACCTTCTTCC 993
Db |||
Qy 660 CCTCACCGAAGCATGTTCAAGCCCGAGTCCGGGTCAAGGACAGCAGTAACTGTGCTGCC 719
Db |||
Qy 994 AGGCCACGGCGGATGTGATGGACCGTTTGGATGGCATGCTCCCGCGCGGATGGT 1047
Db |||
Qy 720 TGGGCACGGCGGCTTCTGGACCGCATCGACAGCTGACGCGCGCCCATTCGGT 773
Db |||
```

Search completed: August 17, 2004, 21:10:38
Job time : 886 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 15:45:38 ; Search time 785 Seconds

(without alignments)
7035.232 Million cell updates/sec

Title: US-09-853-641-1

Perfect score: 1300

Sequence: 1 gaagtccttctgcgaagga.....aggaaaccgattccgctgg 1300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

1: geneseq1980s.*

2: geneseq1990s.*

3: geneseq2000s.*

4: geneseq2001as.*

5: geneseq2001bs.*

6: geneseq2002s.*

7: geneseq2003as.*

8: geneseq2003bs.*

9: geneseq2003cs.*

10: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1300	100.0	1300	6	Aa168796 C. glutam
C 2	1300	100.0	349980	5	Aah68531 C glutami
C 3	1300	100.0	349980	5	Aah68530 C glutami
4	978	75.2	978	4	Aaf67792 Coryneb
5	978	75.2	978	4	Aaf67792 Coryneb
6	855	65.8	855	5	Aah67181 C glutami
7	342.2	26.3	876	7	ACA29916 Prokaryot
8	191.8	14.8	918	7	ACA38529 Prokaryot
9	191.8	14.8	921	7	ACA40735 Prokaryot
C 10	191.8	14.8	110000	4	Continuation (32 o
C 11	191.8	14.8	110000	4	Continuation (32 o
12	168.8	13.0	933	7	ACA38122 Prokaryot
13	164	12.6	939	7	ACA39830 Prokaryot
14	123.4	9.5	297	6	ABN75506 Human syn
C 15	117	9.0	258	5	Aah67180 C glutami
C 16	117	9.0	369	7	ACA00655 C. glutam
17	110.6	8.5	1019	4	ACA59898 Propionib
18	110.6	8.5	1019	4	ACA59898 Propionib
19	69.8	5.4	349980	6	ABQ81849 Bifidobac
20	68.4	5.3	813	7	ACA45738 Prokaryot
21	67.2	5.2	813	7	ACA43503 Prokaryot
22	64	4.9	849	7	ACF69725 Photorhab
C 23	64	4.9	110000	7	Continuation (28 o

24	64	4.9	110000	7	ACF65386_1
C 25	62.4	4.8	19547	4	AAS59601_
C 26	62.4	4.8	19547	7	ACF64530 Propionib
27	59	4.5	843	7	ACA53364 Prokaryot
28	58.8	4.5	936	4	AAS54139 Pseudomon
29	58.8	4.5	936	7	ACA42333 Prokaryot
30	58.4	4.5	2017	6	AAL40183 Isoprenoi
31	58.2	4.5	789	7	ACA35433 Prokaryot
32	58.2	4.5	795	7	ACA37071 Prokaryot
33	56.6	4.4	819	7	ACA23571 Prokaryot
34	55.8	4.3	855	7	ACA27424 Prokaryot
35	55.2	4.2	789	7	ACA36374 Prokaryot
36	55.2	4.2	1733	6	ABQ70612 Listeria
37	55.2	4.2	110000	6	Continuation (14 o
38	55	4.2	819	7	ACA26133 Prokaryot
39	54.6	4.2	748	7	ACA48951 Prokaryot
40	54.6	4.2	858	7	ACA51226 Prokaryot
41	53.2	4.1	846	7	ACA38990 Prokaryot
42	53.2	4.1	945	7	ACF70555 Photorhab
43	53.2	4.1	99629	4	Aaf28550 Genomic f
44	53.2	4.1	110000	7	Continuation (37 o
C 45	53.2	4.1	110000	7	Continuation (12 o

ALIGNMENTS

RESULT 1

Aa168796

ID Aa168796 standard; DNA; 1300 BP.

XX XX

AC Aa168796;

DT 22-JAN-2002 (first entry)

XX C. glutamicum ATCC 13032 cdsA DNA.

DE C. glutamicum ATCC 13032 cdsA DNA.

XX Corynebacterium; cdsA; phosphatide-cytidyltransferase;

KW L-lysine production; amino acid production; medicine; animal nutrition;

KW growth rate; ds.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers

FT RBS 187..195

FT FT /*tag= a

FT CDS 200..1093

FT FT /*tag= b

FT FT /product= "cdsA"

XX DE10021828-A1.

XX 08-NOV-2001.

PD 04-MAY-2000; 2000DE-01021828.

XX 04-MAY-2000; 2000DE-01021828.

XX (DEGS) DEGUSSA AG.

PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

XX Nampootheri M, Moeskel B, Pfeifferle W, Eggeling L, Sahn H;

XX WPI; 2002-018672/03.

XX P-PSDB; AAG80226.

XX New mutant corynebacterium, useful for production of amino acids,

XX especially lysine, has increased activity of phosphatide-cytidyl

XX transferase.

XX Claim 9; Page 10-12; 16pp; German.

XX This invention describes a novel genetically modified corynebacterium

CC bacterium (A) in which activity of the cdsA gene, encoding phosphatide-
 CC cytidyl transferase, is increased. (A) are used for fermentative
 CC production of amino acids, especially L-lysine, useful in human medicine,
 CC animal nutrition and pharmaceuticals. Nucleic acid, or fragments, derived
 CC from the cdsA gene are used (i) as primers for polymerase chain reaction
 CC of cdsA genes or (ii) as probes for isolating cDNA and genes that are
 CC highly homologous with cdsA. Overexpression of the cdsA gene results in
 CC increased production of amino acids, also more rapid growth to higher
 CC cell densities. This sequence encodes the Corynebacterium glutamicum ATCC
 CC 13032 phosphatide-cytidyl transferase (cdsA) described in the method
 CC of the invention
 XX
 SQ

Sequence 1300 BP; 286 A; 289 C; 371 G; 354 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1300; DB 6; Length 1300;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGTCGTTGCTCGCAAGAAAGAAAGTGTAGAGTCTAGAGACCTTTATCGCAATG 60
 DB 1 GAAGTCGTTGCTCGCAAGAAAGAAAGTGTAGAGTCTAGAGACCTTTATCGCAATG 60
 QY 61 GCTTAAGAAATACGTCGTTTCAACGTCGATTGGCGGGGAAACGACGCTTTCTTTGCTT 120
 DB 61 GCTTAAGAAATACGTCGTTTCAACGTCGATTGGCGGGGAAACGACGCTTTCTTTGCTT 120
 QY 121 GCAAGAGTGTGGAGAAATTTTTCGAAATGCTGGCACCATCAACAGTGAATGTTA 180
 DB 121 GCAAGAGTGTGGAGAAATTTTTCGAAATGCTGGCACCATCAACAGTGAATGTTA 180
 QY 181 GAACTTCAAGGAGAACCCATGAATGAACGAGGAGCAATCACCGTCCATGAGATGCC 240
 DB 181 GAACTTCAAGGAGAACCCATGAATGAACGAGGAGCAATCACCGTCCATGAGATGCC 240
 QY 241 CAAACCCAAAAATATGCGGGTGCAGATCTCAAAGCTGCCATTTGCTGGGATCGGACT 300
 DB 241 CAAACCCAAAAATATGCGGGTGCAGATCTCAAAGCTGCCATTTGCTGGGATCGGACT 300
 QY 301 GGGGTCTGCTGTTCTTTTGGGATTTGCTTAAGCCATGGGGTGGTACATCCTCGTTC 360
 DB 301 GGGGTCTGCTGTTCTTTTGGGATTTGCTTAAGCCATGGGGTGGTACATCCTCGTTC 360
 QY 361 AGGTTTATGCTGCAGCAACATGGAAGTGGTAGCAGACTTAAAGAAAGCGGCTATCA 420
 DB 361 AGGTTTATGCTGCAGCAACATGGAAGTGGTAGCAGACTTAAAGAAAGCGGCTATCA 420
 QY 421 TTGCGACTGCGGATATGATCATCGCGGTCAGGCAATCATCTGGCTGATGCGCAAT 480
 DB 421 TTGCGACTGCGGATATGATCATCGCGGTCAGGCAATCATCTGGCTGATGCGCAAT 480
 QY 481 TGGCAGATGGGCAATTTGGGCTCTTTTGGCCACTGTTGGTGTGCTGATGATTTCCG 540
 DB 481 TGGCAGATGGGCAATTTGGGCTCTTTTGGCCACTGTTGGTGTGCTGATGATTTCCG 540
 QY 541 AATTTTCTACATGCGCAAGAAAGAAAGCGGCAACTATTTTGAAGGACACCTCTGTTGG 600
 DB 541 AATTTTCTACATGCGCAAGAAAGAAAGCGGCAACTATTTTGAAGGACACCTCTGTTGG 600
 QY 601 CATCTTGTGCTCAGCTGATTTCCATTTCTCGAAGCTTCGCTCGATGCTGCTGAT 660
 DB 601 CATCTTGTGCTCAGCTGATTTCCATTTCTCGAAGCTTCGCTCGATGCTGCTGAT 660
 QY 661 GCAAAACAAATTCATCCCGGGTACATATTTCAATTTGACGTTTATGATGATCGC 720
 DB 661 GCAAAACAAATTCATCCCGGGTACATATTTCAATTTGACGTTTATGATGATCGC 720
 QY 721 ATCGGATGCGGGATATCGCGGGTGTCTTTTGGATCGCACCAATGGCGGCTT 780
 DB 721 ATCGGATGCGGGATATCGCGGGTGTCTTTTGGATCGCACCAATGGCGGCTT 780
 QY 781 GGTGAGTCCGGAAGTCTTGGGAGGCTTTTGGCGGCTCCATTTGCTTAGGATCGGTAC 840
 DB 781 GGTGAGTCCGGAAGTCTTGGGAGGCTTTTGGCGGCTCCATTTGCTTAGGATCGGTAC 840

QY 841 TGTGCACTCAGTGTTCACCTTCTGCTCGATCACCACCTGGTGGTGTGATCTTGG 900
 DB 841 TGTGCACTCAGTGTTCACCTTCTGCTCGATCACCACCTGGTGGTGTGATCTTGG 900
 QY 901 TTGTGCCCTAGTGTGTGGCCACAGTTGGGTGACTTGGTGTGAGTTCAGCTTCAAACGCGA 960
 DB 901 TTGTGCCCTAGTGTGTGGCCACAGTTGGGTGACTTGGTGTGAGTTCAGCTTCAAACGCGA 960
 QY 961 TTGTGGCATCAAGGATATGTGCAACCTTCTTCAGGCCACGCGGATTGATGACCGTTT 1020
 DB 961 TTGTGGCATCAAGGATATGTGCAACCTTCTTCAGGCCACGCGGATTGATGACCGTTT 1020
 QY 1021 GGATGGATGCTCCCGCGCGGATGTGACGTGTTGATCTCTGAGTGTGATCAGCAGCTC 1080
 DB 1021 GGATGGATGCTCCCGCGCGGATGTGACGTGTTGATCTCTGAGTGTGATCAGCAGCTC 1080
 QY 1081 GTATCCGTGTAAAGCTTGGGCCAGCTTTTAAAGTTCAAAAAACTTGAAGGCCCTGAGGTG 1140
 DB 1081 GTATCCGTGTAAAGCTTGGGCCAGCTTTTAAAGTTCAAAAAACTTGAAGGCCCTGAGGTG 1140
 QY 1141 CATAACGCTGTGCACTCAGCGCTTTTGGCTGTCAAAGTCTTAAAGGGCTTTACGATTT 1200
 DB 1141 CATAACGCTGTGCACTCAGCGCTTTTGGCTGTCAAAGTCTTAAAGGGCTTTACGATTT 1200
 QY 1201 TTCTTAACCTGGGACCGTACTCAAACATCGGCACGCGCCACCAAGCGCCATTAATCAATG 1260
 DB 1201 TTCTTAACCTGGGACCGTACTCAAACATCGGCACGCGCCACCAAGCGCCATTAATCAATG 1260
 QY 1261 CACCGGTAAATGGCTGTAGTAAACCGATTCCGCGCTGG 1300
 DB 1261 CACCGGTAAATGGCTGTAGTAAACCGATTCCGCGCTGG 1300

RESULT 2

AAH68531/c
 ID AAH68531 standard; DNA; 349980 BP.

XX AC AAH68531;

XX DT 26-SEP-2001 (first entry)

XX DE C glutamicum coding sequence fragment SEQ ID NO: 7066.

XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX KW organic acid synthesis; ds.

XX OS Corynebacterium glutamicum.

XX PN EP1108790-A2.

XX PD 20-JUN-2001.

XX PF 18-DEC-2000; 2000EP-00127688.

XX PR 16-DEC-1999; 99JP-00377484.

XX PR 07-APR-2000; 2000JP-00159162.

XX PR 03-AUG-2000; 2000JP-00280988.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX DR WPI; 2001-376931/40.

XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying

XX PT mutation point of a gene, measuring expression of a gene, analyzing

XX PT expression profile or pattern of a gene and identifying homologous gene.

XX PS Disclosure; SEQ ID NO 7066; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium, and identifying a homologue of a gene derived from
CC corynebacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 349980 BP: 80724 A; 98367 C; 90490 G; 80399 T; 0 U; 0 Other;
Query Match 100.0%; Score 1300; DB 5; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAAGTCCTGCTGCAGGAAAGAACTGATGAGGCTTAGAGACCTTTATCGCAATG 60
DB 37324 GAAGTCCTGCTGCAGGAAAGAACTGATGAGGCTTAGAGACCTTTATCGCAATG 37265
QY 61 GCTTAAGAAATACGCTCTTCAACGTCGATTGGCGGGGAAAGCAAGCTTTCTTTTGCTT 120
DB 37264 GCTTAAGAAATACGCTCTTCAACGTCGATTGGCGGGGAAAGCAAGCTTTCTTTTGCTT 37205
QY 121 GCAAGAGTGTTGGAAGAAATTTTTCGAAAATGTTGGCACCATCAACAGTGACATTGTTA 180
DB 37204 GCAAGAGTGTTGGAAGAAATTTTTCGAAAATGTTGGCACCATCAACAGTGACATTGTTA 37145
QY 181 GAAACTTCAAGGAGAACCCATGATGAACGGAGCAACATCACCGGTCCATGAGGATGCC 240
DB 37144 GAAACTTCAAGGAGAACCCATGATGAACGGAGCAACATCACCGGTCCATGAGGATGCC 37085
QY 241 CAACCCCAAAATTAATGCGGTCGAGATCTCAAAAGCTGCCATTCCTGTGGGATCGCACT 300
DB 37084 CAACCCCAAAATTAATGCGGTCGAGATCTCAAAAGCTGCCATTCCTGTGGGATCGCACT 37025
QY 301 GGGGGTCCTGTTCTTTTGGGGATGTCCTTAAGCCCATGGGTTGGTACATCCTCGTTGC 360
DB 37024 GGGGGTCCTGTTCTTTTGGGGATGTCCTTAAGCCCATGGGTTGGTACATCCTCGTTGC 36965
QY 361 AGGTTTATGCTGCAGCAACATGGAAATGTTGTAGCAGACTTAAAGAGCGGCTATCA 420
DB 36964 AGGTTTATGCTGCAGCAACATGGAAATGTTGTAGCAGACTTAAAGAGCGGCTATCA 36905
QY 421 TTGCGCACTGCGGATATGATCATCGCGGTGAGGCAATCATCTGGCTGTCATGCCATT 480
DB 36904 TTGCGCACTGCGGATATGATCATCGCGGTGAGGCAATCATCTGGCTGTCATGCCATT 36845
QY 481 TGGCAGGATGGCATTTTGGCGTCTTTTGTGCGCACTGTGTTGTGCTGATGATTTCCG 540
DB 36844 TGGCAGGATGGCATTTTGGCGTCTTTTGTGCGCACTGTGTTGTGCTGATGATTTCCG 36785
QY 541 AATTTTCTACATGTCAGCGGAAAGAGAGCGGCACTATTGTAGGGACACCTCTGTGGG 600
DB 36784 AATTTTCTACATGTCAGCGGAAAGAGAGCGGCACTATTGTAGGGACACCTCTGTGGG 36725
QY 601 CATCTCTGTGCTCACTGATTCATTTGTTGGAGCTTCGTCGATGCTGCTGAT 660
DB 36724 CATCTCTGTGCTCACTGATTCATTTGTTGGAGCTTCGTCGATGCTGCTGAT 36665
QY 661 GCAAAACAAATTCATCCCGGGTACATATTTTCAATTTGACGTTTCATGCTGTGTGTGATCGC 720
DB 36664 GCAAAACAAATTCATCCCGGGTACATATTTTCAATTTGACGTTTCATGCTGTGTGTGATCGC 36605
QY 721 ATCGGATGTCGGGGATATATCGGGGTGTGTTCTTTGATGATCGCACCCAAATGGCGCCGTT 780
DB 36604 ATCGGATGTCGGGGATATATCGGGGTGTGTTCTTTGATGATCGCACCCAAATGGCGCCGTT 36545
QY 781 GGTGAGTCGAGAGAGCTTTGGAGAGGCTTTCGCGGCTCCATGCTTAGGATCGGTAC 840
DB 36544 GGTGAGTCGAGAGAGCTTTGGAGAGGCTTTCGCGGCTCCATGCTTAGGATCGGTAC 36485

QY 841 TGGTGCACCTCAGTGTTCATCTTCCTGCTCGATCACCACCTGGTGGATGGTGTGATCTTGGG 900
DB 36484 TGGTGCACCTCAGTGTTCATCTTCCTGCTCGATCACCACCTGGTGGATGGTGTGATCTTGGG 36425
QY 901 TTGTGCCCTAGTGTGTGTCGCCACAGTTTGGGTGACCTTGGTTCAGTCGCAGTTTCAACCGCA 960
DB 36424 TTGTGCCCTAGTGTGTGTCGCCACAGTTTGGGTGACCTTGGTTCAGTCGCAGTTTCAACCGCA 36365
QY 961 TTTGGGCAATCAAGGATATGTGCAACCTTCTTCAGGCCACCGCGGATTTGATGACCGCTTT 1020
DB 36364 TTTGGGCAATCAAGGATATGTGCAACCTTCTTCAGGCCACCGCGGATTTGATGACCGCTTT 36305
QY 1021 GGATGGCATGCTCCCGCGCGGATGTCGACGTGGTGTGATCCTGAGTGTGATCAGCAGCTC 1080
DB 36304 GGATGGCATGCTCCCGCGCGGATGTCGACGTGGTGTGATCCTGAGTGTGATCAGCAGCTC 36245
QY 1081 GTATCGCTGTAAAGCTTTGGGCCAGCTTTAAAGTTCAAAAAATTTGAAAGGCGCTGAGGTG 1140
DB 36244 GTATCGCTGTAAAGCTTTGGGCCAGCTTTAAAGTTCAAAAAATTTGAAAGGCGCTGAGGTG 36185
QY 1141 CATACGTTGTGACCTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACCGATT 1200
DB 36184 CATACGTTGTGACCTCAGCGCTTTTGGCTGTCAAAAGTTTAAAGGCTTTTACCGATT 36125
QY 1201 TTCTTAACCTGGCGACGGTACTCAAAACATGCGCACCGCCACCAACAGCGCCATTAATCAATG 1260
DB 36124 TTCTTAACCTGGCGACGGTACTCAAAACATGCGCACCGCCACCAACAGCGCCATTAATCAATG 36065
QY 1261 CACCGGTAATGCTGCTAGTAGGAAACCGATTCCCGCTGG 1300
DB 36064 CACCGGTAATGCTGCTAGTAGGAAACCGATTCCCGCTGG 36025
RESULT 3
AAH68530/c
ID AAH68530 standard; DNA; 349980 BP.
XX
AC AAH68530;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 7065.
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX
KW organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN BP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
XX
PR 07-APR-2000; 2000JP-00159162.
XX
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
PS
PS Disclosure; SEQ ID NO 7065; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC

QY	841	TGGTCACTCAGTGTTCAC	TTCCTCTGCTCGATCACC	ACTGGTGGATGGGTCTGATCTTGG	900
Db	336484	TGGTCACTCAGTGTTCAC	TTCCTCTGCTCGATCACC	ACTGGTGGATGGGTCTGATCTTGG	336425
QY	901	TTTGCCCTAGTTGTGCGC	CAGTTGGGTGACTTGGTT	TGAGTCGCAGTTTCAAACGCGA	960
Db	336424	TTTGCCCTAGTTGTGCGC	CAGTTGGGTGACTTGGTT	TGAGTCGCAGTTTCAAACGCGA	336365
QY	961	TTTGGGCATCAAGGATATG	TCGAACCTTCTCCAGGCCA	CGGCGGATTCATGGACCGTTT	1020
Db	336364	TTTGGGCATCAAGGATATG	TCGAACCTTCTCCAGGCCA	CGGCGGATTCATGGACCGTTT	336305

QY	1021	GGATGGCATGCTCCCGCCGGATGGTACGTGGTTGATCCTCAGTGTGATCAGCAGCTC	1080
Db	336304	GGATGGCATGCTCCCGCCGGATGGTACGTGGTTGATCCTCAGTGTGATCAGCAGCTC	336245
QY	1081	GTATCCGTCGTAAGCTTGGCCAGCTTTAAGTTCAAAAACCTTGAAGCGCTGAGGTG	1140
Db	336244	GTATCCGTCGTAAGCTTGGCCAGCTTTAAGTTCAAAAACCTTGAAGCGCTGAGGTG	336185
QY	1141	CATAACGTGTGCACCTCAGCGCCCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACCGATT	1200
Db	336184	CATAACGTGTGCACCTCAGCGCCCTTTTGGCTGTCAAAGTTTAAAGGGCTTTACCGATT	336125
QY	1201	TTCTTAACCTGCGACGGTACTCAACATGCGCACGCCACCAACAAGCGCCATAATCAATG	1260
Db	336124	TTCTTAACCTGCGACGGTACTCAACATGCGCACGCCACCAACAAGCGCCATAATCAATG	336065
QY	1261	CACCGGTAATGGCTGTAGTAGGAACCGATTCCGGCTGG	1300
Db	336064	CACCGGTAATGGCTGTAGTAGGAACCGATTCCGGCTGG	336025
RESULT 4			
AAF67792			
XX	ID AAF67792 standard; DNA; 978 BP.		
XX	AAF67792;		
XX			
DT	11-APR-2001 (first entry)		
XX	Corynebacterium glutamicum MCT protein encoding DNA SEQ ID NO:99.		
DE	Corynebacterium glutamicum; brevbacterium lactofermentum; MCT;		
XX	membrane construction and membrane transport protein; petroleum spill;		
KW	hydrocarbon degradation; gram positive aerobic bacterium; marker;		
KW	identification; microorganism; fine chemical production; transformation;		
KW	genome mapping; genetic engineering; ds.		
XX			
OS	Corynebacterium glutamicum.		
XX			
PN	WO200100805-A2.		
XX			
PD	04-JAN-2001.		
XX			
PF	23-JUN-2000; 2000WO-IB000926.		
XX			
PPR	25-JUN-1999; 99US-0141031P.		
PR	08-JUL-1999; 99DE-01031454.		
PR	08-JUL-1999; 99DE-01031478.		
PR	08-JUL-1999; 99DE-01031563.		
PR	09-JUL-1999; 99DE-01032122.		
PR	09-JUL-1999; 99DE-01032124.		
PR	09-JUL-1999; 99DE-01032125.		
PR	09-JUL-1999; 99DE-01032128.		
PR	09-JUL-1999; 99DE-01032180.		
PR	09-JUL-1999; 99DE-01032182.		
PR	09-JUL-1999; 99DE-01032190.		
PR	09-JUL-1999; 99DE-01032191.		
PR	09-JUL-1999; 99DE-01032209.		
PR	09-JUL-1999; 99DE-01032212.		
PR	09-JUL-1999; 99DE-01032227.		
PR			

PR	09-JUL-1999;	99DE-01032228.
PR	09-JUL-1999;	99DE-01032228.
PR	09-JUL-1999;	99DE-01032230.
PR	14-JUL-1999;	99DE-01032927.
PR	14-JUL-1999;	99DE-01033005.
PR	14-JUL-1999;	99DE-01033006.
PR	27-AUG-1999;	99DE-01040765.
PR	27-AUG-1999;	99DE-01040766.
PR	27-AUG-1999;	99DE-01040830.
PR	27-AUG-1999;	99DE-01040831.
PR	27-AUG-1999;	99DE-01040832.
PR	27-AUG-1999;	99DE-01040833.
PR	31-AUG-1999;	99DE-01041378.
PR	31-AUG-1999;	99DE-01041379.
PR	31-AUG-1999;	99DE-01041395.
PR	03-SEP-1999;	99DE-01042077.
PR	03-SEP-1999;	99DE-01042078.
PR	03-SEP-1999;	99DE-01042079.
PR	03-SEP-1999;	99DE-01042088.
XX		
PA	(BADI) BASF AG.	

XX
PI Domestius M. Kroeger B. Schroeder H. Zelder O. Haberhauer G.

XX
DR WPI; 2001-071486/08.
DR P-PSDB: AAB76559.

XX Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation.

XX
PS Claim 3: Page 322-323: 1119pp: English: English:

AAAF67743 to AAFA68080 encode the *Corynebacterium glutamicum* membrane construction and membrane transport (MCT) proteins given in AAB76510 to AAB76647. The MCT nucleic acids and proteins are useful in the identification of microorganisms which can be used to produce fine chemicals, for modulating fine chemical production in *C. glutamicum* or related bacteria (e.g. *Brevibacterium lactofermentum*), the typing or identification of *C. glutamicum* or related bacteria, as reference points for mapping *C. glutamicum* genome, and as markers for transformation. AAFA68082 and AAFA68082 represent sequencing primers which are used in an example from the present invention

Sequence 978 BP: 200 A: 220 C: 285 G: 273 T: 0 U: 0 Other; XX

Query Match	75.2%	Score 978;	DB 4;	Length 978;
Best Local Similarity	100.0%	Pred. No. 3.1e-300;		
Matches 978:	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

QY	136	AGAAATTTTTCGAAATGCTGGCACCATCAACAGATGACATCTGTTAGAAACCTTCAGGAGA	195
DB	137	AGAAATTTTTCGAAATGCTGGCACCATCAACAGATGACATCTGTTAGAAACCTTCAGGAGA	60
QY	196	ACCCATGAATGAACCGGAGCAACATCACCGGTCCATAGGATGCCAAACCCAAAATAA	255
DB	197	ACCCATGAATGAACCGGAGCAACATCACCGGTCCATAGGATGCCAAACCCAAAATAA	120
QY	256	TGCGGGTCAGATCTCAAGCTGCCANTTCTCTGGGGATCGGACCTGGGGTCTCTGTTCT	315
DB	257	TGCGGGTCAGATCTCAAGCTGCCANTTCTCTGGGGATCGGACCTGGGGTCTCTGTTCT	180
QY	316	TTTGGGGATTGTCCTAAGGCCCATGGGTGGTACATCCTCGTTGCAAGTTTTATGGCTGC	375
DB	317	TTTGGGGATTGTCCTAAGGCCCATGGGTGGTACATCCTCGTTGCAAGTTTTATGGCTGC	240
QY	376	AGCAACATGGGAAGTTGGTAGCAGACTTAAGAAAGCGGCTATCTTGCACATGCCGAT	435
DB	377	AGCAACATGGGAAGTTGGTAGCAGACTTAAGAAAGCGGCTATCTTGCACATGCCGAT	300
QY	436	TATGATCATCGCGGGTCAGGCAATCATCTCGCTGTGTCATGCCCAATTTGGCAGATGGGCAT	495

PR 01-JUL-1999; 99DE-01030476.
 PR 02-JUL-1999; 99US-0142101P.
 PR 08-JUL-1999; 99DE-01031415.
 PR 08-JUL-1999; 99DE-01031418.
 PR 08-JUL-1999; 99DE-01031419.
 PR 08-JUL-1999; 99DE-01031420.
 PR 08-JUL-1999; 99DE-01031424.
 PR 08-JUL-1999; 99DE-01031428.
 PR 08-JUL-1999; 99DE-01031434.
 PR 08-JUL-1999; 99DE-01031435.
 PR 08-JUL-1999; 99DE-01031443.
 PR 08-JUL-1999; 99DE-01031453.
 PR 08-JUL-1999; 99DE-01031457.
 PR 08-JUL-1999; 99DE-01031465.
 PR 08-JUL-1999; 99DE-01031478.
 PR 08-JUL-1999; 99DE-01031510.
 PR 08-JUL-1999; 99DE-01031541.
 PR 08-JUL-1999; 99DE-01031573.
 PR 08-JUL-1999; 99DE-01031592.
 PR 08-JUL-1999; 99DE-01031632.
 PR 08-JUL-1999; 99DE-01031634.
 PR 08-JUL-1999; 99DE-01031636.
 PR 08-JUL-1999; 99DE-01032125.
 PR 09-JUL-1999; 99DE-01032126.
 PR 09-JUL-1999; 99DE-01032130.
 PR 09-JUL-1999; 99DE-01032186.
 PR 09-JUL-1999; 99DE-01032206.
 PR 09-JUL-1999; 99DE-01032227.
 PR 09-JUL-1999; 99DE-01032228.
 PR 09-JUL-1999; 99DE-01032229.
 PR 09-JUL-1999; 99DE-01032230.
 PR 14-JUL-1999; 99DE-01032922.
 PR 14-JUL-1999; 99DE-01032926.
 PR 14-JUL-1999; 99DE-01032928.
 PR 14-JUL-1999; 99DE-01033004.
 PR 14-JUL-1999; 99DE-01033005.
 PR 14-JUL-1999; 99DE-01033006.
 PR 12-AUG-1999; 99US-0148613P.
 PR 27-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99DE-01040765.
 PR 27-AUG-1999; 99DE-01040766.
 PR 31-AUG-1999; 99DE-01040832.
 PR 31-AUG-1999; 99DE-01041378.
 PR 31-AUG-1999; 99DE-01041379.
 PR 31-AUG-1999; 99DE-01041380.
 PR 31-AUG-1999; 99DE-01041394.
 PR 31-AUG-1999; 99DE-01041396.
 PR 03-SEP-1999; 99DE-01042076.
 PR 03-SEP-1999; 99DE-01042077.
 PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042086.
 PR 03-SEP-1999; 99DE-01042087.
 PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042124.
 PR 03-SEP-1999; 99DE-01042129.
 PR 09-MAR-2000; 2000US-0187970P.
 (BADI) BASF AG.

XX PA

XX PI Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;

XX DR WPI; 2001-137957/14.

XX DR P-PSDB; AAB80192.

XX Nucleic acids from *Corynebacterium glutamicum* encoding metabolic pathway proteins, useful for producing fine chemicals in microorganisms, including organic acids, nonproteinogenic amino acids, and pyrimidine bases.

XX Claim 3; Page 1674; 1737pp; English.

XX AAF71753 to AAF72330 encode the *Corynebacterium glutamicum* metabolic

CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
 CC nucleic acids are useful for the production of fine chemicals in
 CC microorganisms, including organic acids, nonproteinogenic amino acids,
 CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
 CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
 CC vitamins, cofactors, polyketides and enzymes
 XX

SQ Sequence 978 BP; 200 A; 220 C; 285 G; 273 T; 0 U; 0 Other;

Query Match 75.2%; Score 978; DB 4; Length 978;

Best Local Similarity 100.0%; Pred. No. 3.1e-300;

Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 AGAATTTTTCGAAATGCTGCACCATCAACAGTCAGATCTTTAGAACTTCACAGGAGA 195
 Db 1 AGAATTTTTCGAAATGCTGCACCATCAACAGTCAGATCTTTAGAACTTCACAGGAGA 60
 QY 196 ACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCCAACCCCAAAATAA 255
 Db 61 ACCCATGAATGAACCGGAGCAACATCACCGTCCATGAGGATGCCAACCCCAAAATAA 120
 QY 256 TCGGGTTCAGATCTCAAAAGCTGCGATTGCTGCGGATCGGACTGGGGTCTCTGGTCT 315
 Db 121 TCGGGTTCAGATCTCAAAAGCTGCGATTGCTGCGGATCGGACTGGGGTCTCTGGTCT 180
 QY 316 TTTGGGATTGTCTTAAGCCCATGGGTTGGTACATCTCTGTCAGGTTTATGCTGTC 375
 Db 181 TTTGGGATTGTCTTAAGCCCATGGGTTGGTACATCTCTGTCAGGTTTATGCTGTC 240
 QY 376 AGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGCGCTATCATTTGCCATCCGAT 435
 Db 241 AGCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGCGCTATCATTTGCCATCCGAT 300
 QY 436 TATGATCATCGCGGTTCAGGCAATCATCTGCTGTCATGTCGTCATGTCGTCATGTC 495
 Db 301 TATGATCATCGCGGTTCAGGCAATCATCTGCTGTCATGTCGTCATGTCGTCATGTC 360
 QY 496 TTTGGCGTCTTTTGTGGCACTGTGTTGTCGTGTCATGTCATGTCATGTCATGTCATG 555
 Db 361 TTTGGCGTCTTTTGTGGCACTGTGTTGTCGTGTCATGTCATGTCATGTCATGTCATG 420
 QY 556 CACGAAAAAGAGCCCGCAACTATTGAGGACACCTCTGTGGGCATCTTCGTGCTCAC 615
 Db 421 CACGAAAAAGAGCCCGCAACTATTGAGGACACCTCTGTGGGCATCTTCGTGCTCAC 480
 QY 616 CTGGATTCCATTGTTGGGAAGCTTCGCTGCGATGTCGTCGTCATGTCGTCATGTCATG 675
 Db 481 CTGGATTCCATTGTTGGGAAGCTTCGCTGCGATGTCGTCGTCATGTCGTCATGTCATG 540
 QY 676 CCGGGTACATATTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCAT 735
 Db 541 CCGGGTACATATTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCAT 600
 QY 736 GTATATCGCGGTGTGTTCTTTTGGATCGCACCCCAATGGCGCGTGTGTCAGTCAGGAA 795
 Db 601 GTATATCGCGGTGTGTTCTTTTGGATCGCACCCCAATGGCGCGTGTGTCAGTCAGGAA 660
 QY 796 GTCTTGGGAAGGCTTTGCCGCTCCATTGCTTAGGATCGGTCACCTGTCACCTCAGTGT 855
 Db 661 GTCTTGGGAAGGCTTTGCCGCTCCATTGCTTAGGATCGGTCACCTGTCACCTCAGTGT 720
 QY 856 TCATCTCTGTCGATCATCACCTGTCGATGTCGTCATGTCGTCATGTCGTCATGTCGTCAT 915
 Db 721 TCATCTCTGTCGATCATCACCTGTCGATGTCGTCATGTCGTCATGTCGTCATGTCGTCAT 780
 QY 916 GTGGCCACAGTTGGTGACATTTGGTGTAGTCGCACTTCAAGCGGATTTGGGCATCAAGGA 975
 Db 781 GTGGCCACAGTTGGTGACATTTGGTGTAGTCGCACTTCAAGCGGATTTGGGCATCAAGGA 840
 QY 976 TATCTCGAACCTTCTCCAGCCACGGGATTCATGACCGTTCGATGCGATGCTGCTCC 1035
 Db 841 TATCTCGAACCTTCTCCAGCCACGGGATTCATGACCGTTCGATGCGATGCTGCTCC 900

QY 1036 GCGCGGATGTCAGCTGGTTGATCCTGAGTGTGATCAGCAGCTCGTATCCCGTCGTAAAG 1095
Db 901 GCGCGGATGTCAGCTGGTTGATCCTGAGTGTGATCAGCAGCTCGTATCCCGTCGTAAAG 960
QY 1096 CTTGGGCCAGCTTTAAAGT 1113
Db 961 CTTGGGCCAGCTTTAAAGT 978
RESULT 6
AAH67181
ID AAH67181 standard; DNA; 855 BP.
AC
XX AAH67181;
AC
XX 26-SEP-2001 (first entry)
DT XX
DE XX
XX C glutamicum coding sequence fragment SEQ ID NO: 2216.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
KW
XX Corynebacterium glutamicum.
OS
XX EPI108790-A2.
PN
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
PF
XX 16-DEC-1999; 99JP-00377484.
PR
XX 07-APR-2000; 2000JP-00159162.
PR
XX 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
PI
XX WPI: 2001-376931/40.
XX P-PSDB; AAG91962.
DR
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 8; SEQ ID NO 2216; 246pp + Sequence Listing; English.
PS
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids.
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
XX Sequence 855 BP; 159 A; 194 C; 257 G; 245 T; 0 U; 0 Other;
SQ
Query Match 65.8%; Score 855; DB 5; Length 855;
Best Local Similarity 100.0%; Pred. No. 4.2e-261;
Matches 855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 ATGCCCAACCCAAATAATCGGGTCCGAGATCTCAAGCTGCCATTGCTGTGGGATC 295
Db 1 ATGCCCAACCCAAATAATCGGGTCCGAGATCTCAAGCTGCCATTGCTGTGGGATC 60
QY 296 GGACTGGGGTCCCTGTTCTTTTGGGATGTGTCCTAAAGCCCATGGGTGTGATACCTC 355

Db 61 GGACTGGGGTCTGGTCTTTTGGGGAATGCTCTAAGCCCATGGGGTGTGATACCTC 120
QY 356 GTTGACAGTTTTATGGCTGTCAGCAACATGGAAGTTGTTAGCAGACTTAAAGAAGGGCC 415
Db 121 GTTGACAGTTTTATGGCTGTCAGCAACATGGAAGTTGTTAGCAGACTTAAAGAAGGGCC 180
QY 416 TATCAATTTGCCACTGCCGATTAATGATCATCGGGGTAGGCAATCATCTGCTGTCAATG 475
Db 181 TATCAATTTGCCACTGCCGATTAATGATCATCGGGGTAGGCAATCATCTGCTGTCAATG 240
QY 476 CCATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGGCCACTGTGTTGGTCTGATGAT 535
Db 241 CCATTTGGCAGATGGGCAATTTTGGCGTCTTTTGGGCCACTGTGTTGGTCTGATGAT 300
QY 536 TTCCGAATTTTCAATGGCAACGAAAAAAGAGCCCGCAACTATTTCAGGACACCTCT 595
Db 301 TTCCGAATTTTCAATGGCAACGAAAAAAGAGCCCGCAACTATTTCAGGACACCTCT 360
QY 596 GTGGGCATCTTCGCTGCTCACCTGGATTCATTTGCGAAGCTTCGTCGCGATGCTGCG 655
Db 361 GTGGGCATCTTCGCTGCTCACCTGGATTCATTTGCGAAGCTTCGTCGCGATGCTGCG 420
QY 656 CTGATGCAAAAACAATTCATCCCGGTACATATTTTCAATTTGACCTTCATGCTGTGTG 715
Db 421 CTGATGCAAAAACAATTCATCCCGGTACATATTTTCAATTTGACCTTCATGCTGTGTG 480
QY 716 ATCGCATCGATGTGGGGGGTATATCGGGGTGTGTTCTTTGGATCGCACCAATGGCG 775
Db 481 ATCGCATCGATGTGGGGGGTATATCGGGGTGTGTTCTTTGGATCGCACCAATGGCG 540
QY 776 CCGTTGGTGAGTCCGAAAGTCTTTGGGAAGGCTTTGCGGCTCCATTGTCTTAGGATCG 835
Db 541 CCGTTGGTGAGTCCGAAAGTCTTTGGGAAGGCTTTGCGGCTCCATTGTCTTAGGATCG 600
QY 836 GTCACTGGTGCATCAGTGTTCATCTTCGCTCGATCACCACTGGTGGATGGGTGTGATC 895
Db 601 GTCACTGGTGCATCAGTGTTCATCTTCGCTCGATCACCACTGGTGGATGGGTGTGATC 660
QY 896 TTGGGTGTGCGCTAGTGTGTCGCGCAAGTGGGTGACTTGTGAGTCGCGAGTTCAAA 955
Db 661 TTGGGTGTGCGCTAGTGTGTCGCGCAAGTGGGTGACTTGTGAGTCGCGAGTTCAAA 720
QY 956 CGCGATTTGGGATCAAGGATATGTGAAACCTTCTTCCAGGCCACGGGATGATGGAC 1015
Db 721 CGCGATTTGGGATCAAGGATATGTGAAACCTTCTTCCAGGCCACGGGATGATGGAC 780
QY 1016 CGTTGGATGGCATGCTCCCGCGCGATGCTGAGTGTGATCTGAGTGTGATCAGC 1075
Db 781 CGTTGGATGGCATGCTCCCGCGCGATGCTGAGTGTGATCTGAGTGTGATCAGC 840
QY 1076 AGCTCGTATCCGTCG 1090
Db 841 AGCTCGTATCCGTCG 855
RESULT 7
ACA29916
ID ACA29916 standard; DNA; 876 BP.
XX AC
XX ACA29916;
XX
DT 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #11573.
DE
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
XX Corynebacterium diphtheriae.
XX OS
XX WO200271183-A2.
PN
XX 03-OCT-2002.
PD

XX PF 21-MAR-2002; 2002WO-US009107.
 XX PF 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX P-PSDB; ABU26046.
 DR WPI: 2003-029926/02.
 XX DR P-PSDB; ABU26046.
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 14; SEQ ID NO 17786; 1766pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *X. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: the sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 876 BP; 171 A; 184 C; 244 G; 277 T; 0 U; 0 Other;
 Query Match 25.3%; Score 342.2; DB 7; Length 876;
 Best Local Similarity 63.1%; Pred. No. 8.1e-98;
 Matches 527; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
 QY 237 TGCCCAACCCAAAATAATCGGGTCGACATCTCAAAGCTGCCATTGCTGTGGGATCG 296
 DB 32 TGCTAAACCCGAGAAATCGGCTGTGTGAACCTGAAGCCGCGATCAGCGTTGGAAATG 91
 QY 297 GACTGGGGGTCTCGTCTTTTGGGAGTTCCTTAAGCCCATCGGGTGTGATACCTCG 356
 DB 92 GCTTGGGCGCCCTCGTATTCGCTATCTTTGTTATTCCTTTTGGTGTGATCCATTAG 151
 QY 357 TTGCAGGTTTTATGGCTGACGACACATGGAAGTTGGTAGCAGACTTAAAGAGGGCGGT 416
 DB 152 TGCTATCGCAATCGCACTGGGAGGTAGAGCGCCGTTTAAATAGAGCCGGTT 211
 QY 417 ATCATTTGCCACTGCCGATATATGATCATCGGGGTGAGGCAATCATCTGCTGTCTATGCG 476
 DB 212 ATCTTTTGCAACGTTGGGTCAATGCTTTATTTGGGGGCAAGTCAATGTTGTGGCTTAGTTGGC 271
 QY 477 CATTTGGCAGATGGGCATTTTGGCGCTCTTTTGTGCCACTGTGTTGGTGTGCTGATGATT 536
 DB 272 CATTTGGTCCGAGGATAGTGGCTGGTGTGCTGGCGTTGTACTCTCGTAGATGTTG 331
 QY 537 TCCGAATTTTATCAATGGCAGGAAAAAGAGCCCGCAACTATTATTGAGGACACCTCTG 596
 DB 332 GTCCGGCTTTTCCATCATGCTGCTTATGCCACCCGAAAAAATCTATCTGCAGATACTGCCG 391
 QY 597 TGGGCATCTTCGTCGCTACCTGATTCATTCGTAAGCTTCGTAAGCTTCGTCGATGCTGCG 656
 DB 392 TCGCAATTTTGTGTTTAACTGATCCCACTTTTGGTAGTTTTCGGGCAATGCTTTGCG 451
 QY 657 TGATGCAAAACAATTCATCCCGGGTACATATTTTACATGCTTTCATGCTGTGTGTA 716
 DB 452 TGTTTGAACAGAAACAGCACCCGGAAGTACTTTCATGTCACATTCATGTTGTGCGTTA 511
 QY 717 TCGCATCGGATGCGGGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATGGGCGC 776
 DB 512 TCGCCTCTGACGTAGCGGTTTATATTGACGGCTGATGTCGGATCGCACCCCATGGCAC 571
 QY 777 CGTTGGTGTAGTCGGAAGTCTTTGGGAAGGCTTTTCCGGGCTCCATTGCTTTAGGATCGG 836
 DB 572 CTGCGTAAAGCCCAAGAAATCGTGGAGGGCTTTATCGGTTCTGTAGTCTTTGSCATGA 631
 QY 837 TCACTGGTCACTAGTGTTCATCTTCCTGCTCGATCACCACTGGTGGATGGGTGATCT 896
 DB 632 TTGTTGGAGCATTCATCTGTGCTTATTTATTAGGCAATCAATGGTGGTGGGATTAGTTC 691
 QY 897 TGGGTGTGCCCTAGTGTGTCGCGCACGTTGGGTGACTTGGTTGAGTCGAGTTCAAAAC 956
 DB 692 TCGGATTTGGACTTGTGATCTGCGGACACTGGGCGATCTTGTGCAATCCAGTTTAAAG 751
 QY 957 GCGATTTGGGCATCAAGGATATGTGCAACCTTCTTCCAGGCCACGGCGGATTTGATGGACC 1016
 DB 752 GGGAACTCGGATCAAGACATGTCGGAATTTTCCCGGCGCATGGTGTCTGATGGATC 811
 QY 1017 GTTTCGATGGGATGCTCCCGCGCGAGTGTGAGTGTGATCTGATCTGATGATGAT 1071
 DB 812 GCGTTGACGCTATGTTGCCATCGGCTATGTTGATGATGATGATGATGAT 866
 RESULT 8
 ACA38529
 ID ACA38529 standard; DNA; 918 BP.
 XX ACA38529;
 AC ACA38529;
 XX DT 19-JUN-2003 (first entry)
 XX DE Prokaryotic essential gene #20186.
 XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX OS Mycobacterium bovis.
 XX WO200277183-A2.
 XX PD 03-OCT-2002.
 PF 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU34659.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 26399; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 918 BP; 128 A; 278 C; 314 G; 198 T; 0 U; 0 Other;
XX
XX Query Match 14.8%; Score 191.8; DB 7; Length 918;
XX Best Local Similarity 54.4%; Pred. No. 6.5e-50;
XX Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;
XX
XX 257 GCGGGTCGAGATCTCAAGTCGCAATGCTGTGGGATCGGACTGGGGTCTGTGTTCTT 316
XX 88 GCGGCGCGTGTATTCGCGCGACGCGATCGTGTGGGCTTCCATAGGCGCTAGTCTCATC 147
XX
XX 317 TTGGGATGTCCTAAGCCATCGGTGTGTACATCTCGTTGCGAGTTTATGGCTGCA 376
XX
XX 148 GCGGTGCTGGTGTTCGTTCCGCGGCTTGGTTGCTTCATCGTGGCGTCGCCACCTTGTC 207
XX
XX 377 GCACATCGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGATT 436
XX
XX 208 GCTACCCATGAGTGTGCGGAGGTTCGCGGAAGCGGCTATCTCATCCCGTTATCCCG 267
XX
XX 437 ATGATCATCGGGTCAGGCAATCATCTGGCTGTATCGGCTATGGCAGCATGGGATT 496
XX
XX 268 TTGCTGATTTGGCGGCGCGCGCGGTGTGGTGTACCTGGCGGCTTCGGCGGCTGGCGCA 327
XX
XX 497 TTGGCGTCTTTTGTGGCCACTGTGTGTGTGTGTATTTCCGAATTTCTACAATGGC 556
XX
XX 328 TTGGCGGCTTTGGTGGCATGTCTGTCTGTCATGATTTGGGACTGTTTCATGACGAC 387
XX
XX 557 ACGGAAAAAGAACGCC-----GCAACTATTAGGGACACC 592

388 AGCGTGAAGCCCGACGACGCGCGGTGACCCGTCGCGGGAACACTACTTGTCCGACGTC 447
593 TCTGTGGGCATCTTCGTCCTCACCTGGATTCCATGTTGGGAAGCTTCGCTGCGATCGTG 652
448 TCGGCCACGGTCTTCTTGGCGGTGTGGGTCCCATGTTCTCTCTTTCGCGGCAATCGTG 507
653 TCGGTGATGCAAAACAATTCATCCCGGTACATATTTTTCATTTTTCGCTTTCATGCTGTGT 712
508 GTCATCCCGGAAATGGCTCGGATGGG-----TGTTCTGCATGATGATCGCG 555
713 GTGATCGCATCGGATGTGGCGGGTATATCGCGGTGTGTTCCTTTTGGATCGCACCCCAATG 772
556 GTCATCGCTTCCGATGTGCGCGCTACGCCGTGGGGTGTGTTTGGCAAGCATCCGATG 615
773 GCGCGGTGTGTAGTCCGAAAGTCTTGGAAAGCTTTGCGGCTCCATGTTCTTAGGA 832
616 GTTCCGACGATCAGCCCAAGAGTGTGGGAGGCTTTGCGGTTCTGCTGTGTGCGGG 675
833 TCGGTCACTGTGTGCACTCAGTGTTCACCTCTCTGTCTCGATCACCACCTGTGTGATGGTGTG 892
676 ATCACCAGCAAGCATCATCACCGGACTTCTGTGCGCAAAACGCGGTGGATTGGTGA 735
893 ATCTTGGGTGTGCGCTAGTGTGTGTGCGCCACGTTGGGTGACTGTGGTGGTGGTGGTGGT 952
736 CTGCTCGCGGTCTTTTCTGT 795
953 AAACGCGATTTGGGATCAAGATATGTCGAACCTTCTTCCAGGCGACGCGGATTTGATG 1012
796 AAACGTGACCTCGGATCAAGATATGTCGAACCTTCTTCCAGGCGACGCGGATTTGATG 855
1013 GACCGT 1071
856 GACCGGT 914

RESULT 9
ACA40735
ID ACA40735 standard; DNA; 921 BP.
AC ACA40735;
XX
XX 19-JUN-2003 (first entry)
DE Prokaryotic essential gene #22392.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Mycobacterium tuberculosis.
XX WC200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 08-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX P-PSDB; ABU36865.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.


```
WP AAI99682_43 4300001 4410000
WP AAI99682_44 4400001 4411529

Query Match
Best Local Similarity 14.8%; Score 191.8; DB 4; Length 110000;
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;

QY 257 GCGGTCGAGATCTCAAGCTGCATTCGTGGGATCGGATCGGGGTCCTGGTTCTT 316
Db 91531 GCCGCCGTGATTGCGCGCAGCGATCGTGTGGCGCTTCCATAGCCCTAGTCTCATC 91472
QY 317 TTGGGGATTGCTTAAGCCATAGGGGTGTGTACATCTCTCTTCAGGTATTATGGCTGCA 376
Db 91471 GCGGTGCTGTGTTCGTTCCGCGGTTTGGGTTCGCCATCTGCGCGCTCGCCACTTGGTC 91412
QY 377 GCACATGGAAGTTGTAGCAGACTTAAAGAGCGCGCTATCATTTGCCACTGCCGATT 436
Db 91411 GCTACCCATGAGTGTGGGAGGTTCGGGAGCGGCTATCTCATCCGGTTATCCCG 91352
QY 437 ATGATCATCGCGCTCAGGCAATCATCTGGCTGTCATCGGCAATTTGGCAGATGGGCAAT 496
Db 91351 TTGCTGATTGGCGGCAAGCGCGGTGTGGCTGACCTGGCGGCTTCGGCGCGTCCGCGCA 91292
QY 497 TTGGGCTTTTGGCCACATCTGTTGCTGTGATTTTCGAAATTTCTACAATGGC 556
Db 91291 TTGGCGGCTTTGTTGGCATGCTGTTGCTGTGATTTTCGAAATTTCTACAATGGC 91232
QY 557 ACGGAAAAAAGAGCC-----GCAACTATTTAGGGACACC 592
Db 91231 AGCGTGACGCCCGCGACGACCGCGGTCACCGTCGCGGGAACACTACTTGTCCGAGTC 91172
QY 593 TCTGTGGGCACTCTGCTCCTCAGTATTCATTTGTCGGAAGCTTCGTCGATGCTG 652
Db 91171 TCGGCGACGCTCTCTTGGCGGCTGGGTGCCATTTCTGCTCTTTTCGGCGCAATGCTG 91112
QY 653 TCGCTGATGCAAAACAATTCATCCGCGGTACATATTTCAATTTGAGCTTCATGCTGT 712
Db 91111 GTCTACCGGAAATGCTTCGGATGG-----TGTTCTGATCATGATCGCG 91064
QY 713 GTGATCGCATCGGATGTCGGCGGATATTCGCGGTGTTCTTTTGGATGCCACCCATG 772
Db 91063 GTGATCGCTTCGATGTCGGCGGTACCGCTGGGGTGTCTTTTGGCAATCCGATG 91004
QY 773 GCGCGTGTGTGAGTCCGAAGCTTCGGAAGCTTTGCGGCTCCATGTTCTTAGGA 832
Db 91003 GTTCCGACGATCAGCCCGAAGTCGTGGAGGCTTTGCGGTTCTGCTGTGTGCGG 90944
QY 833 TCGGTCACTGTGTCACTCAGTGTTCACCTTCTGTCTGATCACCACCTGTGGATGGGTG 892
Db 90943 ATCACCAGCAAGATCATCACCGGACTTTCTGTGTCGGCAAAACGCCGTGATTTGTGCA 90884
QY 893 ATCTTGGGTTGTGCCCTAGTTGTGCGCCAGTTCGGTGGGTGCTGTTGAGTCCGAGTTC 952
Db 90883 CTGCTCGGCTGCTTTTCTGCTCACCACCGGCTGGGCGACCTGTGGATCGGAGTTC 90824
QY 953 AAACCGGATTTCGGCATCAAGGATATGCAACCTTTCTTCCAGGCCACCGGGGATTTGATG 1012
Db 90823 AAAGTGACCTCGGATCAAGACATGGCCGCTGCTACCGGCCACCGGGTCTGATG 90764
QY 1013 GACCGTTTGGATGGATGCTCCCGCGGATGGTGTGATGCTGATGCTGATG 1071
Db 90763 GACCGCTCGAGGATGATGCTTCCGCGTGGCGGCTGATAGTCTCTCACACTGCT 90705
```

RESULT 11
AAI99683_31/c
Continuation (32 of 44) of AAI99683 from base 3100001 (Mycobacterium tuberculosis strain
WP Sequence split into 44 fragments LOCUS AAI99683 Accession AAI99683
WP Fragment Name Begin End
WP AAI99683_00 1 110000
WP AAI99683_01 100001 210000
WP AAI99683_02 200001 310000
WP AAI99683_03 300001 410000

```
AAI99683_04 400001 510000
AAI99683_05 500001 610000
AAI99683_06 600001 710000
AAI99683_07 700001 810000
AAI99683_08 800001 910000
AAI99683_09 900001 1010000
AAI99683_10 1000001 1110000
AAI99683_11 1100001 1210000
AAI99683_12 1200001 1310000
AAI99683_13 1300001 1410000
AAI99683_14 1400001 1510000
AAI99683_15 1500001 1610000
AAI99683_16 1600001 1710000
AAI99683_17 1700001 1810000
AAI99683_18 1800001 1910000
AAI99683_19 1900001 2010000
AAI99683_20 2000001 2110000
AAI99683_21 2100001 2210000
AAI99683_22 2200001 2310000
AAI99683_23 2300001 2410000
AAI99683_24 2400001 2510000
AAI99683_25 2500001 2610000
AAI99683_26 2600001 2710000
AAI99683_27 2700001 2810000
AAI99683_28 2800001 2910000
AAI99683_29 2900001 3010000
AAI99683_30 3000001 3110000
AAI99683_31 3100001 3210000
AAI99683_32 3200001 3310000
AAI99683_33 3300001 3410000
AAI99683_34 3400001 3510000
AAI99683_35 3500001 3610000
AAI99683_36 3600001 3710000
AAI99683_37 3700001 3810000
AAI99683_38 3800001 3910000
AAI99683_39 3900001 4010000
AAI99683_40 4000001 4110000
AAI99683_41 4100001 4210000
AAI99683_42 4200001 4310000
AAI99683_43 4300001 4403765

Query Match
Best Local Similarity 14.8%; Score 191.8; DB 4; Length 110000;
Matches 456; Conservative 0; Mismatches 347; Indels 36; Gaps 2;

QY 257 GCGGTCGAGATCTCAAGCTGCATTCGTGGGATCGGATCGGGGTCCTGGTTCTT 316
Db 85640 GCCGCCGTGATTTCGGCGAGCGATCGTGTGGCGCTTCCATAGGCTAGTCTCATC 85581
QY 317 TTGGGGATTGCTTAAGCCATAGGGGTGTGTACATCTCTCTTCAGGTATTATGGCTGCA 376
Db 85580 GCGGTGCTGTGTTCGTTCCGCGCTTTGGGTTGCCATCGTGGCGCTTCGCGATC 85521
QY 377 GCACATGGAAGTTGTAGCAGACTTAAAGAGCGCGCTATCATTTGCCACTGCCGATT 436
Db 85520 GCTACCCATGAGTGTGCGGAGGTTCGGGAGCGGCTATCTCATCCGGTTATCCCG 85461
QY 437 ATGATCATCGCGCTCAGGCAATCATCTGGCTGTCAATGGCCATTTGGCAGATGGGCAAT 496
Db 85460 TTGCTGATTGGCGGAGCGCGGTGTGGCTGACCTGGCGGCTTCGGCGCGCTCGGCGCA 85401
QY 497 TTGGGCTTTTGGCCACATCTGTTGCTGTGATTTTCGAAATTTCTACAATGGC 556
Db 85400 TTGGCGGCTTTGGTGCAATGCTGTGCTGATTTTGGCGAGCTGTTCTATGCAAGAC 85341
QY 557 ACGGAAAAAAGAGCC-----GCAACTATTTAGGGACACC 592
Db 85340 AGCGTGACGCCCGCGACGACCGCGGTCGACCGTTCGGCGGAAACTACTTGTCCGAGTC 85281
QY 593 TCTGTGGGCACTCTGCTCCTCAGTATTCATTTGTCGGAAGCTTCGTCGATGCTGCTG 652
Db 85280 TCGGCCACGCTCTCTTCCCTGGCGGTGTGGTCCCATTTGTTCTCTTTTCGGCGCAATGCTG 85221
```

QY 653 TCCTGATGCAAAACAATTCCATCCCGGTACATATTTTCATTTGACGTTTCATGCTGTGT 712
Db 85220 GTCTACCCGGAATAAGCTCGGATGGG-----TGTTCTGATGATGATCGCG 85173
QY 713 GTATCGCATCGGATGCGCGGATATATCGCGGGTGTCTTCTTTGGATCGCACCAATG 772
Db 85172 GTATCGCTTCGATGTCGGCGGTACGCGGTGGGGTGTCTTTGGCAAGATCCGATG 85113
QY 773 GCGCCGTTGTGATGTCGAAGAAGCTTTGGGAAGCTTTGCGGTTCCATGTTCTTAGGA 832
Db 85112 GTTCCACGATCAGCCCGAAGAAGCTTGTGGAGGCTTTGCGGTTCTGCTGTGCGGG 85053
QY 833 TCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 892
Db 85052 ATCACCGCAACATATCACCGGACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 84993
QY 893 ATCTGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952
Db 84992 CTGCTCGGCTGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 84933
QY 953 AAACGGATTTGGGATCAGGATATGTCGAACCTTCTTCCAGGCCACGGCGGATGATG 1012
Db 84932 AAACGGATTTGGGATCAGGATATGTCGAACCTTCTTCCAGGCCACGGCGGATGATG 84873
QY 1013 GACCGTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1071
Db 84872 GACCGGCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 84814

RESULT 12

ACA38122
ID ACA38122 standard; DNA; 933 BP.

AC ACA38122;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #19779.

KW Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

OS Mycobacterium avium.

PN WO20027183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Walli D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI: 2003-029926/02.

XX P-PSDB; ABU34252.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 25992; 1766pp; English.

CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

QY Sequence 933 BP; 102 A; 318 C; 343 G; 170 T; 0 U; 0 Other;

Query Match 13.0%; Score 168.8; DB 7; Length 933;

Best Local Similarity 52.8%; Pred. No. 1.4e-42;

Matches 439; Conservative 0; Mismatches 357; Indels 36; Gaps 2;

QY 257 GCGGTCGAGATCTCAAAGCTGCCATTGCTGTGGGGATCGGACTGGGGTCTGTTCTT 316
Db 103 GCGGACGTGACCTCGGGGCGCCCATCGGGTGGCGCGGTATCGGCGCGTCTATC 162
QY 317 TTGGGGATGTCCTAAGCCATGCGGTTGGTATACCTCTGTTGACAGTTTATGGCTGCA 376
Db 163 GTACGCTGGTGTTCGCGCCCGCTTCTGGGTGCGGATCGTCGCGATGCCCATCTCGT 222
QY 377 GCAACATGGGAAGTTGGTAGCAGACTTAAAGAAGCGCGCTATCTTCGCACTGCCGATT 436
Db 223 GCGACCATGAGTGTGTCGCGCGCTGCGGAGCGGATAGCTATCCCGTATTCCG 282
QY 437 ATGATCATCGCGGTCAGCAATCATCTGGGTGTCTATGCCCATTTGGCAGTGGCAATT 496
Db 283 TTGTCGCGCGCGCGCAGCTACGGTGTGGCTGACCTGGCGGTTCCACGCGCGCGCG 342
QY 497 TTGGCGTCTTTTGTGCGCACCTGTGTTGGTGTGATGATTTCCGAATTTCTACATGCG 556
Db 343 TTGGCGCGCTTCGGCGTACCGGTGTGCGCTGCTGTTCTGGCGGTGTTTCTATGAGAC 402
QY 557 ACGGAAAAAGAACGCCGC-----AACTATTTGAGGACACC 592
Db 403 AACCCAAAGCGCCCGAGCCGTTGCGCGGTTGCGCCTCGCGGAACCTACCTCGCGACGG 462
QY 593 TCTGTGGGCACTTCGCTGCTCAGCTGGATTCATGTTCCGAACTGTCGATGCGATGCTG 652
Db 463 TCGGCCACCGTCTTCTGCGCTGCTGCGGTGCTGTTGTCGCTCTCTTGGCGCGCTGCTG 522
QY 653 TCGTGTATGCAAAACAATTCCATCCCGGTACATATTTTCATTTGAGGTTTCATGCTGT 712
Db 523 -----GTCTATCCCGCGACGGTTCGCGCGGGGTGTTCTGCCGTATGATCAC 570
QY 713 GTATCGCATCGGATGCGCGGTATATCGCGGTGTGTTCTTTGGATCGCACCAATG 772
Db 571 GTGTGGCGCTTCGACGTGCGCGGTACCGGTGGCGGTGCTGTGTTTGGCAAGCATCCGATG 630
QY 773 GCGCGGTGTTGAGTTCGGAAGACTCTTGGGAAGCTTTGCGGCTCCATTGTTCTTAGGA 832

Db 631 GTCCGCGGATCAGCCCAAGAAATCTCTGGAGGCGCTGGCGGCTCGTGTGTGGGT 690
 QY 833 TCGGTCACTGGTGCATCACTAGTTTCACTTCCTGCTCGATCACCATGCTGATGGGTGTG 892
 Db 691 ATCACCAGCGGACACCTCGCGGCGACTTCTTCTGCGCGGAAGCGCCGCTGGGCGCG 750
 QY 893 ATCTTGGGTGTGTCCTAGTTGTGTGCGCCAGCTTGGGTGACTTGGTTGAGTTCGCGAGTTC 952
 Db 751 CTGCTGGGGGTGTGCTGTGTTCACTGSCACCTCGCGGACCTGTGTGAGTTCGCGAGTC 810
 QY 953 AAACGGGATTTGGGATCAAGGATATGCAACCTTCTTCAAGGCCAGCGGATTTGATG 1012
 Db 811 AACGGGACTGGGATCAAGGATATGCAACCTTCTTCAAGGCCAGCGGATTTGATG 870
 QY 1013 GACCGTTTGGATGGCATGCTCCCGCGCGGATGCGTGTGATGCTGCA 1064
 Db 871 GACCGGCTGGAGGTGTGCTGCTGCGGTGCGCGGCTGAGCGTGTCTGA 922

RESULT 13

ACA39830
 ID ACA39830 standard; DNA; 939 BP.

XX ACA39830;

XX ACA39830;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #21487.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 drug design; gene.

XX Mycobacterium leprae.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-VAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU35960.

XX New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 27700; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing of a gene in an operon required for
 proliferation or the activity of the gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 939 BP; 144 A; 248 C; 310 G; 237 T; 0 U; 0 Other;

Query Match 12.6%; Score 164; DB 7; Length 939;
 Best Local Similarity 52.6%; Pred. No. 4.7e-41;

Matches 440; Conservative 0; Mismatches 375; Indels 21; Gaps 3;

QY 257 GCGGTCGAGATCTCAAAGCTGCCATTCCTGTGGGATCGGACTGGGGTCTCTGTTCTT 316
 Db 97 GCGGACGTAATCTCCCGCGCGGATCGCGTGGGCTTAAGTATGGTGTGTTCTCGTC 156
 QY 317 TTGGGGATGTCTTAAGCCCATGCGGTGTGTATCATCTCTGTTGAGGTTTATGCTGCA 376
 Db 157 GCAACGCTGGTGTGCTCCGGAATCTGGTGTCTTGTGCGCTGGCGCATTTTCTT 216
 QY 377 GCAACATGGGAAGTTGGTAGCAGACTTAAAGAGCGGCTATCATTTGCCACTGCCGAT 436
 Db 217 GCTAGCCATGAGGTGTGCGCGGCTACGGAAGCTGATATGTGATTCGGGCTATCCG 276
 QY 437 ATGATCATCGCGGCTCAGGCAATCATCTGGCTGTTCATGGCCATTTGGCAGATGGCATT 496
 Db 277 CTGCTCATCGTGGGCACTTCAACGCTCTGGTTAACTTGGCGGTATCGCACGTTGGTGA 336
 QY 497 TTGGGCTCTTTTGGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 550
 Db 337 TTGGGCTGT 396
 QY 551 AATGGCACGAAAAAGAGCGCGCAACTATTATTGAGGACACTCTGTGGGCATCTTCTGTG 610
 Db 397 AACACAGACGACGCAATCTCGGAGACTTGGCAGTCCGCGGTGTCAATTAATCTG 456
 QY 611 -----CTACCTGGATTCATTTGGGAAGCTTCTGTGTGATCTGTGTGTGTGTGTGTGT 665
 Db 457 CCGGACGCGTCTCGGCACCGCTTTTGTGGCGGCTGGGTCTCGCTGTGTGTGTGTGTGT 516
 QY 666 ACA-----ATTCCATCCCGGTACATATTTCATTTTGGAGCTTCACTGTGTGTGTG 715
 Db 517 GATTTGTGTGTGTATCCGAGAGATGGTCCGCGCGGTATTTCTCCCTGATGATTCGCGTG 576
 QY 716 ATGCGCATCGGATGTGGCGGATATATCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 775
 Db 577 GTAGCTTCGATGTGGGTGTGTACACAGTAGGGGTACTTTTTCGGCAAAACATCCACTGTT 636
 QY 776 CGTTTGT 835
 Db 637 CCAAGGATTTAGCCCGATTAAGTCTGTGGAGGATTCGCGGTGTGTGTGTGTGTGTGTGT 696
 QY 836 GTCACTGTGTGCACTGATGTTTCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
 Db 697 ACCGCAACCATCTGACTGTACTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 756
 QY 896 TTGGGT 955
 Db 757 CTCAGTTTGT 816

QY 956 CGCGATTGGGATCAAGGATATGTCGAACCTTCTTCAGGCCACGGCGGATGATGAC 1015
 Db 817 CGCGACTCGGCATCAAGATATGGCCCGCTGCTACCGCGCACGGTGGCTGATGAC 876
 QY 1016 CGTTTGGATGGATGCTCCCGGCGGATGCTGACCTGTTGATCTGAGTGTGAT 1071
 Db 877 CGGCTTGATGCTGCTGCGCTCGGCGGTGTAGTTGGACCATCTTACGTTGCT 932

RESULT 14
 ID ABN75506
 XX ABN75506 standard; cDNA; 297 BP.
 AC ABN75506;
 DT 08-JUL-2002 (first entry)
 XX
 DE Human synthase-like ORF453 cDNA, SEQ ID NO:905.
 KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neutropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiact; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 PD
 PF 29-NOV-2001.
 PF
 XX 24-MAY-2001; 2001WO-US017076.
 PR
 XX 24-MAY-2000; 2000US-0206690P.
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shimkets RA;
 XX
 DR WPI; 2002-106200/14.
 XX P-PSDB; ABP31480.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 XX transplantation.
 PS Claim 1; Page 495; 2508pp; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,

CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases
 XX

SQ Sequence 297 BP; 51 A; 91 C; 102 G; 53 T; 0 U; 0 Other;

Query Match 9.5%; Score 123.4; DB 6; Length 297;
 Best Local Similarity 68.0%; Pred. No. 2.1e-28;
 Matches 172; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 854 GTTCACCTTCTCTCTCATCACCATGCTGGTGGATGGTGTGATCTTGGTTGTCCTAGTT 913
 Db 1 GTGCACTTCTCTCATCACCATGCTGGTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGT 60
 QY 914 GTGTGGCCACGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 973
 Db 61 CTCTGGCCACCATGGGACCTCTGTGAGACGAGTTCACGCGGAGCTGGGAATCAAG 120
 QY 974 GATATGTGAACTTCTTCCAGGCCACGGCGGATGATGACCGTTTGGATGGCATGCTC 1033
 Db 121 GACATGTGAACTTGTCTCCCGCCACGGCGGCTGATGGACCGCTTGGACGGGATGCTG 180
 QY 1034 CCGGCCGCGATGCTGACGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1093
 Db 181 CCGGCCGCGACGGCGACCTACATGCTGTCAACGCCGGCGGATCATGAACCCGCTGTAG 240
 QY 1094 AGCTTGGGCCAGC 1106
 Db 241 CCGGCCCTGGCAGC 253

RESULT 15
 AAH67180/c
 ID AAH67180 standard; DNA; 258 BP.
 XX
 AC AAH67180;
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 2215.
 XX
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 OS Corynebacterium glutamicum.
 PN EP1108790-A2.
 XX
 XX 20-JUN-2001.
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 93JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI: 2001-376931/40.
DR P-PSDB; AAG91961.
XX

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX

Claim 8; SEQ ID NO 221-5; 246pp + Sequence Listing: English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

XX SQ Sequence 258 BP; 57 A; 69 C; 61 G; 71 T; 0 U; 0 Other;

Query Match 9.0%; Score 117; DB 5; Length 258;
Best Local Similarity 100.0%; Pred.No. 2.le-26;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1184 AAGGCGTTTACGGATTCTTTCTTAACCTGGCGACGGTACTCAAAACATGGCGACGCCACCAAC 1243
Db 258 AAGGCGTTTACGGATTCTTTCTTAACCTGGCGACGGTACTCAAAACATGGCGACGCCACCAAC 199
Qy 1244 AAGCGCCATAATCAATGACCGGTAATGGCTGTAGTAGGAACCGATTCCGGCTGG 1300
Db 198 AAGCGCCATAATCAATGACCGGTAATGGCTGTAGTAGGAACCGATTCCGGCTGG 142

Search completed: August 17, 2004, 17:09:36
Job time : 793 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 16:43:55 ; Search time 5204 Seconds

(without alignments)

7459.807 Million cell updates/sec

Title: US-09-853-641-1

Perfect score: 1300

Sequence: 1 gaagctcttcttcgaagga.....aggaaaccgattccgggtgg 1300

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*

2: em_estlin:*

3: em_estlin:*

4: em_estlin:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_estl:*

11: gb_estl:*

12: gb_estl:*

13: gb_estl:*

14: gb_estl:*

15: em_estfun:*

16: em_estfun:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158.4	12.2	526	28	BH740535
2	56.8	4.4	885	13	BX425603
3	51.6	4.0	712	13	BX416727
4	48.4	3.7	648	14	CB287945

5	48.4	3.7	695	14	CB287953
6	45.8	3.5	627	13	BQ560202
7	45.8	3.5	629	13	BY718485
8	45.8	3.5	649	12	BM934705
9	45.8	3.5	779	14	CB289962
10	45.8	3.5	1322	11	BC026505
11	45.8	3.5	1386	29	AY399407
12	45.6	3.5	829	28	BH402821
13	45.2	3.5	435	9	AI844295
14	45	3.5	1201	13	BX381961
15	44.8	3.4	403	12	BU8894397
16	44.8	3.4	614	12	BJ168290
17	44.8	3.4	632	12	BJ602826
18	44.8	3.4	638	12	BJ579676
19	44	3.4	379	28	B67568
20	43.4	3.3	755	12	BJ610177
21	43	3.3	768	14	CF880314
22	43	3.3	823	14	CB907156
23	42.8	3.3	1201	13	EX361080
24	42.6	3.3	812	14	CK235970
25	42.2	3.3	471	12	BJ579141
26	42	3.2	884	29	CNS006U0
27	42	3.2	895	29	CNS0071A
28	41.8	3.2	558	14	CD871811
29	40.8	3.1	1101	29	CNS00L72
30	40.6	3.1	534	13	EX522779
31	40.4	3.1	1201	13	EX355644
32	40.2	3.1	612	13	BU808466
33	40.2	3.1	699	13	EX424825
34	39.8	3.1	698	29	CG897864
35	39.8	3.1	705	29	CG898009
36	39.4	3.0	870	29	AGC49346
37	39.4	3.0	1201	13	EX356664
38	39.2	3.0	546	28	BH900070
39	39	3.0	627	12	BM934188
40	39	3.0	849	13	EX462111
41	39	3.0	1101	29	CNS016HG
42	38.8	3.0	534	12	BM129973
43	38.8	3.0	686	13	BM041691
44	38.6	3.0	583	9	AV545973
45	38.6	3.0	703	9	AV783262

ALIGNMENTS

RESULT 1
BH740535
LOCUS

DEFINITION

BH740535 526 bp DNA linear GSS 22-FEB-2002
cpbav00114 Corynebacterium pseudotuberculosis Bacterial Artificial
Chromosomes library Corynebacterium pseudotuberculosis genomic
clone CB2.4-f10r, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH740535 1 GI:18861038

GSS.

Corynebacterium pseudotuberculosis

Corynebacterium pseudotuberculosis

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1 (bases 1 to 526)

Azevedo, V., Fachin, M.S. and Oliveira, S.

Construction and Characterization of Corynebacterium

pseudotuberculosis Bacterial Artificial Chromosomes library

Unpublished (2002)

Contact: Azevedo V

Department of General Biology

Federal University of Minas Gerais

Av. Antonio Carlos, 6627, Campus Pampulha, Belo Horizonte, Minas

Gerais, Brazil

Tel: 00 55 31 3499 2610

Fax: 00 55 31 3499 2610

Email: vasco@mono.icb.ufmg.br

SEQUENCES WERE BASECALLED WITH PHRED AND VECTOR WAS MASKED WITH

CROSSMATCH. SEQUENCES WERE THEN TRIMMED FROM BOTH ENDS TO REMOVE LOW QUALITY BASES AND MASKED VECTOR.

Seq primer: M13 Reverse

Class: BAC ends.

Location/Qualifiers

1..526

/organism="Corynebacterium pseudotuberculosis"

/mol_type="genomic DNA"

/strain="1002"

/db_xref="taxon:1719"

/clone="Cb2.4-f10r"

/lab_host="Escherichia coli DH10B"

/clone_lib="Corynebacterium pseudotuberculosis Bacterial

Artificial Chromosomes library"

/note="Vector: pBelobAC11; Genomic DNA was digested with

HindIII and randomly cloned into HindIII digested and

dephosphorylated pBelobAC11."

FEATURES

source

Query Match 12.2%; Score 158.4; DB 28; Length 526;

Best Local Similarity 66.6%; Pred. No. 2.2e-35; Mismatches 138; Indels 5; Gaps 4;

Matches 285; Conservative 0;

634 AAGCTTCGCTGCATGCTGCTGCTGATGCAAAACAATTCATCCCGGGTACATATTTTCAT 693

1 AAGCTTCGAGTAAGCTCTCTCAGTTCAAAATGACGTAGCTCTCTGGGCAATTTTGT 60

694 TTGACGTTTCATGCTGTGTGATCGATCGATCGATGCGGGGTATATCGGGGTGTGT 753

61 TGTAACTTCATGCTTTGCGTGTGTTCTCTGACGTTTGGTGTGATGCTGCTGTGTAT 120

754 CTTTCGATCGCAACCAATGGCGCGTGTGAGTCCGAGAAAGTCTTGGGAAGCTTTGC 813

121 GTTTGATCCCAACCGATGGCTCTGAGTGGTCCAAAGAAATCGTGGGAAGTTTCGC 180

814 CGGCTCCATTTCTTAGGATCGGTCATCGTGGTGCATCACTAGTGTTCATCTCTGCTCATCA 873

181 GGGTTTCAGTTATTTTGGAATGACTGTGCGGCTATTGAGTGTGATTTCTCTGGAATTA 240

874 CCACGTGGTGGTGTGATCTGGTGTGCTAGTGTGCGGCTAGTGTGCGGCTAGTGTGGA 933

241 CACCTGGTGGTGGGCTTAAATCTAGGCTTGGATTTGGCTTACCTTAGCGGA 300

934 CTTGTGTGA-GTCCGAGTTCAAAACCGATTTGGGCATCAAGGATATGTGCAACCTTC-TT 991

301 TTTAGTGGANGTCAAGTTCAGCGAGAACTAGGTATTAAAGACATGTCCGCTATGCNTT 360

992 CCAGGCCACGGCGG--ATTGATGACCGTTTGGATGGCATGCT-CCCGGCCGCGATGGT 1048

361 CCGGCCACGGTGGTCTTATTGGATAGCTCCGACGGAATGTGCCCTTCTGCCATGTT 420

1049 ACGTGGTT 1056

421 ACGTGGT 428

RESULT 2

EX425603

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX425603 885 bp mRNA linear EST 15-MAY-2003

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID: CL0BB022ZA07FPl.

FEATURES

source

1..885

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CL0BB022ZA07"

/tissue_type="NEUROBLASTOMA"

/clone_lib="Homo sapiens NEUROBLASTOMA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

ORIGIN

Query Match 4.4%; Score 56.8; DB 13; Length 885;

Best Local Similarity 3.3%; Pred. No. 5.2e-05;

Matches 16; Conservative 161; Mismatches 312; Indels 0; Gaps 0;

338 TGGGGTTGTACATCTCGTTCAGGTTTATGGCTGCAGCAACATGGGAAGTTGGTAGC 397

339 TNNKKKKKNTTKTKTKKKKKKKTKKKKKKKNNNNNNNNNNNNNNNNNNNNNN 398

398 AGACTTAAAGAGGGGGTATCATTTGCCACTGCGCATATGATCATCGGGGTGAGCA 457

399 NNANNN 458

458 ATCATCTGCTGCTGATGGCCATTTCGACGATGGGCAATTTTGGCGCTCTTTTGGCCACT 517

459 KKKKKKKKKKKKKKKKKKKKKKKKKKKNNNNNNNNNNNNNNNNNNNNNNNNNN 518

518 GTGTTGCTGCTGATTTTCGAAATTTTCAATATGGCACGGAAGAAAGAGCCGCAAC 577

519 KKKKKKKKKKKKKKKKKKKKKKKKKKKNNNNNNNNNNNNNNNNNNNNNNNNNN 578

578 TATTGAGGACACCTCTCTGCGGCTCTCGTCTCACCTGATTCATTTTCGGAAGC 637

579 NNN 638

638 TTGCTGCGATGCTGCTGATGCAAAACAATTCATCCCGGGTACATATTTTCATTTG 697

639 KNN 698

698 ACGTTCATGCTGCTGATCGCATGATGTTGGGGGGTATATCGGGGTGTTCTTT 757

699 NNN 758

758 GGATCGCACCAATGGCGCGCTTGGTGGTCCGAGAGAGTCTTGGGAAGGCTTTGCGCGC 817

759 KKKKKKKKKKKKKKKKKKKKKKKKKKKNNNNNNNNNNNNNNNNNNNNNNNNNN 818

818 TCCATTGTC 826

819 KNNNNNNNN 827

RESULT 3

EX416727

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EX416727 712 bp mRNA linear EST 15-MAY-2003

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

University of Minnesota
1971 Commonwealth Ave., St. Paul, MN 55108, USA
Tel: 612-625-6735
Fax: 612-625-5203
Email: murta001@umn.edu

Email: murta001@umn.edu

The PHRED quality scores of this sequence are supplied below: 51 51

51	40	51	51	51	40	40	40	37	37	40	39	39	31	31	31	35
35	40	34	23	23	29	29	45	51	56	51	51	51	51	39	39	39
35	35	56	40	40	40	40	40	45	46	42	37	40	40	38	38	38
38	38	40	45	45	43	43	43	43	43	43	43	40	56	56	56	51
51	51	45	45	41	41	46	46	56	56	56	43	38	43	43	43	43
43	43	43	43	43	36	36	36	43	40	40	42	46	46	51	56	56
41	41	45	45	41	41	45	45	45	45	43	43	43	43	44	46	41
41	41	56	56	56	56	56	51	56	56	51	45	45	51	51	51	56
41	41	56	56	56	56	56	51	56	56	51	45	45	51	51	51	56

56 43 43 43 43 43 43 43

[illegible][illegible]

1 1
 2 2
 3 3
 4 4
 5 5
 6 6
 7 7
 8 8
 9 9
 10 10
 11 11
 12 12
 13 13
 14 14
 15 15
 16 16
 17 17
 18 18
 19 19
 20 20
 21 21
 22 22
 23 23
 24 24
 25 25
 26 26
 27 27
 28 28
 29 29
 30 30
 31 31
 32 32
 33 33
 34 34
 35 35
 36 36
 37 37
 38 38
 39 39
 40 40
 41 41
 42 42
 43 43
 44 44
 45 45
 46 46
 47 47
 48 48
 49 49
 50 50
 51 51
 52 52
 53 53
 54 54
 55 55
 56 56
 57 57
 58 58
 59 59
 60 60
 61 61
 62 62
 63 63
 64 64
 65 65
 66 66
 67 67
 68 68
 69 69
 70 70
 71 71
 72 72
 73 73
 74 74
 75 75
 76 76
 77 77
 78 78
 79 79
 80 80
 81 81
 82 82
 83 83
 84 84
 85 85
 86 86
 87 87
 88 88
 89 89
 90 90
 91 91
 92 92
 93 93
 94 94
 95 95
 96 96
 97 97
 98 98
 99 99
 100 100
 101 101
 102 102
 103 103
 104 104
 105 105
 106 106
 107 107
 108 108
 109 109
 110 110
 111 111
 112 112
 113 113
 114 114
 115 115
 116 116
 117 117
 118 118
 119 119
 120 120
 121 121
 122 122
 123 123
 124 124
 125 125
 126 126
 127 127
 128 128
 129 129
 130 130
 131 131
 132 132
 133 133
 134 134
 135 135
 136 136
 137 137
 138 138
 139 139
 140 140
 141 141
 142 142
 143 143
 144 144
 145 145
 146 146
 147 147
 148 148
 149 149
 150 150
 151 151
 152 152
 153 153
 154 154
 155 155
 156 156
 157 157
 158 158
 159 159
 160 160
 161 161
 162 162
 163 163
 164 164
 165 165
 166 166
 167 167
 168 168
 169 169
 170 170
 171 171
 172 172
 173 173
 174 174
 175 175
 176 176
 177 177
 178 178
 179 179
 180 180
 181 181
 182 182
 183 183
 184 184
 185 185
 186 186
 187 187
 188 188
 189 189
 190 190
 191 191
 192 192
 193 193
 194 194
 195 195
 196 196
 197 197
 198 198
 199 199
 200 200
 201 201
 202 202
 203 203
 204 204
 205 205
 206 206
 207 207
 208 208
 209 209
 210 210
 211 211
 212 212
 213 213
 214 214
 215 215
 216 216
 217 217
 218 218
 219 219
 220 220
 221 221
 222 222
 223 223
 224 224
 225 225
 226 226
 227 227
 228 228
 229 229
 230 230
 231 231
 232 232
 233 233
 234 234
 235 235
 236 236
 237 237
 238 238
 239 239
 240 240
 241 241
 242 242
 243 243
 244 244
 245 245
 246 246
 247 247
 248 248
 249 249
 250 250
 251 251
 252 252
 253 253
 254 254
 255 255
 256 256
 257 257
 258 258
 259 259
 260 260
 261 261
 262 262
 263 263
 264 264
 265 265
 266 266
 267 267
 268 268
 269 269
 270 270
 271 271
 272 272
 273 273
 274 274
 275 275
 276 276
 277 277
 278 278
 279 279
 280 280
 281 281
 282 282
 283 283
 284 284
 285 285
 286 286
 287 287
 288 288
 289 289
 290 290
 291 291
 292 292
 293 293
 294 294
 295 295
 296 296
 297 297
 298 298
 299 299
 300 300
 301 301
 302 302
 303 303
 304 304
 305 305
 306 306
 307 307
 308 308
 309 309
 310 310
 311 311
 312 312
 313 313
 314 314
 315 315
 316 316
 317 317
 318 318
 319 319
 320 320
 321 321
 322 322
 323 323
 324 324
 325 325
 326 326
 327 327
 328 328
 329 329
 330 330
 331 331
 332 332
 333 333
 334 334
 335 335
 336 336
 337 337
 338 338
 339 339
 340 340
 341 341
 342 342
 343 343
 344 344
 345 345
 346 346
 347 347
 348 348
 349 349
 350 350
 351 351
 352 352
 353 353
 354 354
 355 355
 356 356
 357 357
 358 358
 359 359

43 43 43 43 43 43 43

51 51 51 51 51 51 51 51

45 45 45 45 45 45 45 45 45 45

43 43 43 37 37 41 45 45

43 43 43 43 51 51 56 56

56 56 56 56 46 46 51

57	46	45	40	40	44	45	45
----	----	----	----	----	----	----	----

30	40	37	40	45	45
H	F	F	F	F	F

33 0F 3C 3C 0F 3C 3C 0F 3C

40 45 45 45 40 35

35 35 40 40 40 40 51 51

```

FEATURES
  high quality sequence scop: 202.
  location/Qualifiers
    1. .648
      /organism="Sus scrofa"
      /mol_type="mRNA"
      /db_xref="taxon:9823"
      /clone="PPSUBLIB_56E12"
      /tissue type="Peyer's patch"
      /clone lib="UMNMPM3"

```

FEATURES source

ORIGIN

277	ATTTTGAACAACAGATTCTTGACATGCGGGATAATGGACAGCTTTGACTGTCAGTAAT	336
1032	TCCCGGCGCGAGTGTGACGTGGTTGATCCTTGAGTGTGATCAGACAGCTCTGATCCGTCGT	1091
337	TGATGGCAACGTTCTGTGACGCTGTACATCATCAGCTTCATCAGGGGTCCGAAATCCCACGA	396
1092	AAAGCTTGGCGCCAGCTTTAAGTTTCAAAAAAATTTGAAAGGCGCTGA	1136
397	AAGTGCTCCAGCAGCTGTTAGTGTCTTCAACAGAACACGACGAGTTAA	441

BY718485 629 bp mRNA linear EST 17-DEC-2002
BY718485 RIKEN full-length enriched, adult male medulla oblongata
Mus musculus cDNA clone 6330520I175', mRNA sequence.

Nikaido, I., Oagi, N., Suzuki, H., Yamana, K.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, I.,
Shimobach, C., Goleburi, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schriml, L. M., Kapin, A., Matsuda, H.,
Batalov, S., Baisel, K. W., Blake, J. A., Bradt, D., Brusica, V.,
Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
Fletcher, C. P., Forrest, A., Frazer, K. S., Gaasterland, S.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gusticoldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Hartshorn, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, P. M., King, B. I., Konagaya, A.,
Koshkin, I. V., Lee, Y., Lehar, S., Lyons, P. A., Maglott, D. R.,
Maltais, P., Marchionni, L., McKenzie, L., Miki, H., Nagasima, T.,
Numata, K., Okido, J., Pavan, W. J., Pertea, G., Pesole, G.,
Petkovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tonita, M.,
Vatardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Cinnici, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Akawaka, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sakaki, D., Shibata, K.,
Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
Rogers, J., Birney, E. and Havashizaki, Y.

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL 22354683
 MEDLINE
 PUBMED 12468851
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Db	402	ATTTTGCAACACGAGTTCCTGGACATTCCTGGCGGATAATGCAGAGTTTGACTCTCAGTATT	461
QY	1032	TCCCGCGCGCATGTGTGACGTCGTGTGATCTCTGAGTGTGATCAGCAGCTCGTATCCGTCGT	1091
Db	462	TGATGGCAACGTTCTGTGCACGTGTACATCACCAGCTTCATCAGGGGTCCGAATCCCGCA	521
QY	1092	AAAGCTTGGCGCAGCTTTAAGTTCACAAAACCTTGAAGGCGCTGA	1136
Db	522	AAAGTCTCCAGCAGCTGTAGTGTCTCAACCAGAACACGAGTTAA	566
RESULT 8			
LOCUS	BM934705		
DEFINITION	UI-M-BH2.2-adq-g-11-0-UI.r1 NIH BMAP M.S3.2 Mus musculus cDNA clone		
ACCESSION	BM934705	649 bp	linear
VERSION	BM934705.1	UI-M-BH2.2-adq-g-11-0-UI 5',	mRNA sequence.
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 649)		
JOURNAL	Bonaldu, M.F., Lennon, G. and Soares, M.B.		
MEDLINE	Normalization and subtraction: two approaches to facilitate gene		
PIRMEF	discovery		
	Genome Res. 6 (9), 791-806 (1996)		
	97044477		
	990549		

8889548
PUBMED
COMMENT

Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 REVERSE.

Location/Qualifiers
1..649
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH2.2-aqk-g-11-01"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP M S3.2"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP M S3.2 library is a subtracted library of a
series, ultimately derived from a mixture of individually
tagged normalized libraries from ten regions of the mouse
brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal

FEATURES
source

gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S3.2, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S3'2) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S2 clones from which 3' ESTs had been derived were used as a driver in a hybridization with the NIH_BMAP_M_S2 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B

bacteria (Life Technologies) to generate the NIH BMAP M 83.2 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"

ORIGIN

Query Match 3.5%; Score 45.8; DB 12; Length 649;
Best Local Similarity 53.3%; Pred. No. 0.08;
Matches 120; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 915 TGTGGCCACGTTGGGTGACTTGGTTGAGTCGCAGTTCAAACCGGATTTGGGCATCAAGG 974
DB 200 TGATGGCCCATTTGGAGGCTCTTTCCAGTGATTCAAAAGAGCTTCAAATCAAGG 259
QY 975 ATATGTGCAACCTTCTTCCAGGCCACGCGGATGTGATGGACCGTTTGGATGGCATGC--- 1031
DB 260 ATTTGCAACACAGATTCCTGACATGCGGGATATGGACAGGTTTGACTGTCAATT 319
QY 1032 TCCGCGCGGAGTGTGACGTGGTTGATCTGTGATGTGATCAGCAGCTCGTATCCGTGCT 1091
DB 320 TGATGGCAACGTTTGGTCACGTTGATCATCACCAGCTTCATCAGGGGTCCGAATCCAGCA 379
QY 1092 AAAGCTTGGGCCAGCTTTAAGTTCAAATAAATTGAAAGGCGCTGA 1136
DB 380 AAGTGTCCAGCAGCTTTAGTGTCTCAACGACAGCAGTTAA 424

RESULT 9
CB289962 779 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-POO-cdv-n-05-0-UI.r1 NIH_BMAP_F00 Mus musculus CDNA clone
DEFINITION IMAGE:6835110 5', mRNA sequence.

ACCESSION CB289962
VERSION CB289962.1 GI:28611377

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 779)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 779

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6835110"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_F00"

/note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I;

Site_2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

FEATURES

source

FEATURES

source

1. .1322
/organism="Mus musculus"
/mol_type="mRNA"

with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAGAGACC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 3.5%; Score 45.8; DB 14; Length 779;
Best Local Similarity 53.3%; Pred. No. 0.09;
Matches 120; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 915 TGTGGCCACGTTGGGTGACTTGGTTGAGTCGCAGTTCAAACCGGATTTGGGCATCAAGG 974
DB 70 TGATGGCCCATTTGGAGGCTCTTTTGGCAGTGGATTCAAAGAGCTTCAAATCAAGG 129
QY 975 ATATGTGCAACCTTCTTCCAGGCCACGCGGATTCATGACCGTTTGGATGGCATGC--- 1031
DB 130 ATTTGCAACACAGATTCCTGACATGCGGGATATGGACAGGTTTGACTGTCAATT 189
QY 1032 TCCGCGCGGAGTGTGACGTGGTTGATCCTGAGTGTGATCAGCAGCTCGTATCCGTGCT 1091
DB 190 TGATGGCAACGTTTGGTCACGTTGATCATCACCAGCTTCATCAGGGGTCCGAATCCAGCA 249
QY 1092 AAAGCTTGGGCCAGCTTTAAGTTCAAATAAATTGAAAGGCGCTGA 1136
DB 250 AAGTGTCCAGCAGCTTTAGTGTCTCAACGACAGCAGTTAA 294

RESULT 10

BC026505

LOCUS

BC026505

DEFINITION

BC026505

ACCESSION

BC026505.1

VERSION

BC026505.1

KEYWORDS

HTC.

SOURCE

Mus musculus (house mouse)

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1322)

Strausberg, R.

Direct Submission

Submitted (02-APR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: http://www-shgc.stanford.edu

Contact: (Dickson, Mark) mcdpaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAK Plate: 44 Row: i Column: 9

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis

This clone has the following problem: frame shifted.

Location/Qualifiers

1. .1322

/organism="Mus musculus"

/mol_type="mRNA"

Query Match 3.5%; Score 45.8; DB 29; Length 1386;
Best Local Similarity 53.3%; Pred. No. 0.13;
Matches 120; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

Query Match 3.5%; Score 45.6; DB 28; Length 829;
Best Local Similarity 62.1%; Pred. No. 0.11;

```

Matches 72; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 926 TTGGTACTGTGTTAGTGCAGTTCAAACGGGATTTGGGCATCAGGATATGTGAAC 985
Db 795 TGGGAGATTGGTGGAGTGCAGCTCAAAGAACTTTCAATGTAAGATTCGGGGAT 736
QY 986 CTTCTTCAGCCACGGCGGATTGATGACCGTTTGATGGCATGCTCCCGGCGC 1041
Db 735 ATAATCCGGGCACGAGGTGATTGGACAGGCTGGATAGCTTATACCTTGGC 680

RESULT 13
A1844295 435 bp mRNA linear EST 15-JUL-1999
LOCUS UI-M-AL1-ahm-d-04-0-UI.s1 NIH BMAP_MCO.N Mus musculus cDNA clone
DEFINITION UI-M-AL1-ahm-d-04-0-UI 3', mRNA sequence.
ACCESSION A1844295
VERSION A1844295.1 GI:5488201
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 435)
COMMENT Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 889548
CONTACT: Chin, H
NATIONAL INSTITUTE OF MENTAL HEALTH
6001 EXECUTIVE BLVD. ROOM 7N-7190, MSC 9643, BETHESDA, MD
20892-9643, USA
TEL: 301 443 1706
FAX: 301 443 9890
EMAIL: mstg@mail.nih.gov
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
available by the means that is soon to be determined. When NIH
determines the means for distribution of the BMAP cDNA clones, this
record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
Location/Qualifiers
1..435
/mol_type="mRNA"
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clones="UI-M-AL1-ahm-d-04-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_MCO_N"
/note="Vector: pUT3D-Pac (Pharmacia) with a modified
polylinker; site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_MCO_N library is a normalized library constructed
from mouse cortex. The tag is a string of 5 nucleotides
present between the Not I site and the oligo-dt track.
The library was constructed as described by Bonaldo,
Lennon and Soares, Genome Research 6: 791-806, 1996.
Tissue provided by Ms. Annie Novakovich, Zivic-Miller
Laboratories.
TAG_TISSUE=prefrontal-cortex
TAG_LIB=NIH_BMAP_MCO_N
TAG_SEQ=GCTCA"

ORIGIN
Query Match 3.5%; Score 45.2; DB 9; Length 435;
Best Local Similarity 53.7%; Pred. No. 0.095;
Matches 117; Conservative 0; Mismatches 98; Indels 3; Gaps 1;

Matches 72; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 922 CACGTTGGTGGTACTTGGTTAGTGCAGTTCAAACGGATTTGGGCATCAAGGATATGTC 981
Db 9 CTCATTGGAGGCTTCTTTGGCAGTGGATTCAAAGAGCTTTCAAATCAAGGATTTGTC 68
QY 982 GAACCTTCTTCAGGCCACCGCGGATTCATGACACCGTTTGGATGCGATGC---TCCCGGC 1038
Db 69 AAACACGATTCTGGACATGGCGGATAATGACACAGGTTTGACTGTGATGATTTGATGCG 128
QY 1039 CCGGATGGTGCAGTGGTTGATCTCAGTGTGATCAGCAGCTGATCGTCTGTAAGCTT 1098
Db 129 AACGTTGGTGCACGTGTATACATCCAGCTTCATCAGGGGTCCGAATCCACGAAAGTCT 188
QY 1099 GGGCCAGCTTTTAAGTTCAAAAACCTTGAAGGCGCTGA 1136
Db 189 CCAGCAGCTGTTAGTGTCTTCAACACGAGACAGCAGTTAA 226

RESULT 14
BX381961/c 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSODI072YF05 3-PRIME, mRNA sequence.
ACCESSION BX381961
VERSION BX381961.1 GI:30453007
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 1201)
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
COMMENT Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CSODI072CC03NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI072YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 3.5%; Score 45; DB 13; Length 1201;
Best Local Similarity 5.3%; Pred. No. 0.2;
Matches 32; Conservative 223; Mismatches 373; Indels 2; Gaps 1;

QY 64 TAAGAATTACGTCGTTTCAACGTCGATTGGCGGGGAAACGACGCTTTCTTTTGTCTGCA 123
Db 1077 KVVYGMEMNKKKMMKVMVMGMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMM 1018
QY 124 AGAGTGTGTTGGAGAAATTTTTCGAAATGCTGGCACCATCAACAGTGCATGTTAGAA 183
Db 1017 KTGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 958
QY 184 ACTTCAAGGAGAACCCATGAATGAACCGGAGCAACATCACCGGTCCATGAGGATGCCCAA 243
Db 957 KVMKMBGVNKKTKMMMBKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMM 898
QY 244 ACCCAAAAATAATGCGGGTCGAGATCTCAAAGTCCATGCCATGCTGTGGGATCGGACTGG 303
Db 244 ACCCAAAAATAATGCGGGTCGAGATCTCAAAGTCCATGCCATGCTGTGGGATCGGACTGG 303

```


C	1	191.8	14.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
C	2	191.8	14.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
C	3	64.2	4.9	1503	4	US-09-252-931A-4643	Sequence 4543, Ap
C	4	64.2	4.9	1854	4	US-09-252-931A-4542	Sequence 4542, Ap
C	5	58.8	4.5	1041	4	US-09-252-931A-14536	Sequence 14536, A
C	6	58.8	4.5	1209	4	US-09-252-931A-14851	Sequence 14851, A
C	7	58.2	4.5	870	4	US-09-489-039A-5219	Sequence 5219, Ap
C	8	54.4	4.2	7060	4	US-09-221-017A-479	Sequence 479, App
C	9	53.2	4.1	891	4	US-09-540-236-198	Sequence 198, App
C	10	53.2	4.1	99629	4	US-09-595-002-37	Sequence 37, Appl
C	11	50.8	3.9	879	4	US-09-543-681A-2323	Sequence 2323, Ap
C	12	45.4	3.5	858	4	US-09-328-352-3131	Sequence 3131, Ap
C	13	44	3.4	1461	4	US-09-252-931A-4311	Sequence 4311, Ap
C	14	43.8	3.4	948	4	US-09-543-681A-4311	Sequence 4311, Ap
C	15	43.8	3.4	1259	2	US-08-672-814D-12	Sequence 2949, Ap
C	16	43.8	3.4	1259	3	US-09-333-696-12	Sequence 12, Appl
C	17	43.4	3.3	16535	4	US-08-961-527-74	Sequence 12, Appl
C	18	41.8	3.2	579	4	US-09-134-000C-2347	Sequence 2347, Ap
C	19	39.8	3.1	2051	1	US-08-672-814D-1	Sequence 1, Appl1
C	20	39.8	3.1	2051	3	US-09-333-696-1	Sequence 1, Appl1
C	21	38	2.9	786	4	US-09-134-001C-2129	Sequence 2129, Ap
C	22	37.6	2.9	2051	4	US-09-282-218A-13	Sequence 1, Appl1
C	23	37.6	2.9	2103	4	US-09-282-218A-13	Sequence 13, Appl
C	24	36.8	2.8	1830121	4	US-09-557-884-1	Sequence 1, Appl1
C	25	36.8	2.8	1830121	4	US-09-643-990A-1	Sequence 1, Appl1
C	26	36	2.8	1419	4	US-09-489-039A-1247	Sequence 1247, Ap
C	27	35.8	2.8	430	4	US-09-621-976-16656	Sequence 16656, A

QY 690 TCATTTTGACGTTTCATGCTGTGTGATCGCATCGATGTGGCGGGGTATATCGCGGGTG 749
Db 1097 TGATCATCCCGTGTGCTGTGCTGTGGCGCGGATATCGGCGCTACTTCTCCGGCA 1156
QY 750 TGTCTTTGGATCGACCCCAATGGCGCGCTTGTGTAGTCCGAAGACTCTTGGGAAGCT 809
Db 1157 AGGCTTTCGGCAAGCGCAAGCTTGGCGCGCGGTTCAGCCCCGCAAGAGCTGGGAAGCG 1216
QY 810 TTGCGGCTCCATGTCTTAGGATCGG---TCACTGGTGTCACTCAGTGTTCATCTTCCTGC 866
Db 1217 TTACGCGCGCTTGGCGGCACGCTGGCGATCACTCGCGGTTCACCGGGCT 1276
QY 867 TCGATCACCACTGGTGATGGGTGTGATCTTGGGTGTGCCCTAGTGTGTGCGGCACGT 926
Db 1277 GGTCCCTCGCGCGCTGTCTTGGCGCTGTGGCGCGCGCTGTGTGTGTCTCGA 1336
QY 927 TG---GGTCACTTGGTGTGATCGCACTTCAAGCGGATTTGGGCATCAAGGATATGTGCA 983
Db 1337 TGTGCGCGACCTTACCGAAGCATGTTCAAGCGCCATCCGATCAAGGACGACGCA 1396
QY 984 ACCTTCTTCCAGGCCACGGCGGATTTGATGACCGTTTGGATGGCATGCTCCCGGCGCA 1043
Db 1397 ACTGTGTCGCGCCACCGGTGGCGTGTGATCGCATCGACAGCCTGACGCGAGCATCC 1456
QY 1044 TGGTG 1048
Db 1457 CGGTG 1461

RESULT 4
US-09-252-991A-4542
; Sequence 4542, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4542
; LENGTH: 1854
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4542

Query Match 4.9%; Score 64.2; DB 4; Length 1854;
Best Local Similarity 52.3%; Pred. No. 6.8e-11;
Matches 191; Conservative 0; Mismatches 168; Indels 6; Gaps 2;
QY 690 TCATTTTGACGTTTCATGCTGTGTGATCGCATCGATGTGGCGGGGTATATCGCGGGTG 749
Db 261 TGATCATCGCGTGTGCTGTGTGGCGCGGATATCGGCGCTACTTCTCCGGCA 320
QY 750 TGTCTTTGGATCGACCCCAATGGCGCGCTTGTGTAGTCCGAAGACTCTTGGGAAGCT 809
Db 321 AGGCTTTCGGCAAGCGCAAGCTTGGCGCGCGGTTCAGCCCCGCAAGAGCTGGGAAGCG 380
QY 810 TTGCGGCTCCATGTCTTAGGATCGG---TCACTGGTGTCACTCAGTGTTCATCTTCCTGC 866
Db 381 TTACGCGCGCTTGGCGGCACGCTGGGATCACTTGGCGGTTCGGCTCTACCGCGCT 440
QY 867 TCGATCACCACTGTGTGATGGGTGTGATCTTGGGTGTGCCCTAGTGTGTGCGCCACGT 926
Db 441 GGTCCCTCGGCGCGCTGTCTTGGCCCTGTCTCGCGCGCGCTGGTGTGTCTCTCGA 500
QY 927 TG---GGTCACTTGGTGTGATCGGCAAGTTCAAAACCGGATTTGGGCATCAAGGATATGTGCA 983

Db 501 TCGTCGCGACCTTACCGAAGCATGTTCAAGCGCCHATTCCGAATCAAGGACAGCAGCA 560
QY 984 ACCTTCTTCCAGGCCACGGCGGATTTGATGACCGTTTGGATGGCATGCTCCCGCGCCGA 1043
Db 561 ACCTGTGCGCGCCACCGTGGCGTGTGATGCGATCGACAGCTTACGCGCAGCCATCC 620
QY 1044 TGGTG 1048
Db 621 CGGTG 625

RESULT 5
US-09-252-991A-14536
; Sequence 14536, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14536
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14536

Query Match 4.5%; Score 58.8; DB 4; Length 1041;
Best Local Similarity 50.9%; Pred. No. 3.1e-09;
Matches 166; Conservative 0; Mismatches 157; Indels 3; Gaps 1;
QY 722 TCGGATGTGGCGGGTATATCGCGGTGTGTTTGGATCGCACCCCAATCGCGCGGTG 781
Db 692 TCCGACGTGATCGATACATCTCGGCAAGCTGTTTCGCAAGCGCAAGATCGTCCCAAC 741
QY 782 GTGATCGGAAGACTCTTGGGAAGCTTTCGGGCTCCATGCTTAGGATCGGTCACT 841
Db 742 CTGTGCGGCTCGAAGACCGTTCGAAGCTTCTGCGCGCATCGCCCTGCCACCCCATC 801
QY 842 GGTGCACTCAGTGTTCACCTTCTGCTCGATCACCACTGCTGGATGGGTGTGATCTTGGGT 901
Db 802 GCGCTGCTGTGTGATCACCCGTTCAACCTTGCAGGCGCGCTTGTATCGCGTG 861
QY 902 TGTGCCCTAGTGTGTGCGCCACGTTGGGTGACTTGGTTGATGTCAGTTCAAACGCGAT 961
Db 862 CTGATCAACCTCTCGGGCTTC---TTCGCGCGCTTGTGATGTGCGCGCATCAAGCGCGAC 918
QY 962 TTGGGCATCAAGGATATGTGCAACCTTCTTCAGCGCCACGGCGGATTTGATGACCGTTG 1021
Db 919 CBTGGGTCAAGGACTTGGGGGCACATGATCGAAGTCACTGCGTGGCATGCTCAGCGCGCTG 978
QY 1022 GATGGCATCTCCCGCGCGCGATGGT 1047
Db 979 GACTCGGTGTCTTCGCGCGCGCGAT 1004

RESULT 6
US-09-252-991A-14831/c
; Sequence 14831, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14831
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14831

Query Match 4.5%; Score 58.8; DB 4; Length 1209;
Best Local Similarity 50.9%; Pred. No. 3.4e-09;
Matches 166; Conservative 0; Mismatches 157; Indels 3; Gaps 1;

QY 722 TCGGATGTGGCGGGTATATCCCGGGTGTGTCTTTGGATCGCACCCCAATGGCGCCGTTG 781
DB 480 TCCGACGTGATGACATCTCGGCAAGCTGTTCGCGAAGCGCAAGATCGCTCCCAAC 421
QY 782 GTGAGTCGGAAGATCTTGGGAAGCTTCCCGCTCCATTGTCTAGATCGGTCACT 841
DB 420 CTGTGCGCGTCGACGATCGAAGGCTTCGTGCGCGGATCGCCCTGGCCACCGCCATC 361
QY 842 GGTGCACTCACTGTCTCTCTCGATCACCACCTGGTGGATGGTGTGATCTTGGT 901
DB 360 GCGCCTCGTGTGGTGAATCACCCGTTCAACCTCTGGCAGCGGCTTGAICGCGCTG 301
QY 902 TGTGCCCTAGTTGTGCGGCACCTTGGGTGACTTGGTGGATGGATGGATGCAACGCGAT 961
DB 300 CTGATCAACCTCTGCGGCTTC--TTGCGCGGCTGTGTGATGTGCGGATCAAGCGGAC 244
QY 962 TTGGCATCAAGGATGTGCAACTTCTCCAGGCGCACGGGATGTGACCGCTTGG 1021
DB 243 CGTGGGTCAAGGACTGGGGCATGATCGAAGTCAAGTCAAGTCAAGTCAAGTCAAG 184
QY 1022 GATGCGATGTCGCGCGCGGATGTT 1047
DB 183 GACTCGGTGTCTTCGCGCGCGGAT 158

RESULT 7

US-09-489-039A-5219
; Sequence 5219, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 5219

; LENGTH: 870

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-5219

Query Match 4.5%; Score 58.2; DB 4; Length 870;
Best Local Similarity 48.8%; Pred. No. 4.4e-09;
Matches 187; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

QY 702 TCATGCTGTGTGATCGCATCGGATGCGGGGTATATCGGGGTGTGTCTTTGGAT 761
DB 482 TCATGATCTGTGTGGGCGGACATCTGGCGCTATATGTTGGTAAATGTTGCA 541
QY 762 CGCACCAATGGCGCGTGTGTGAGTCCGAAGAAGTCTTGGGAAGC---TTTGGCGGT 818
DB 542 AGCACAAGTGGCGGCGAAGGTCTACCGGGTAAACCTTGGCAGGCGTTTGTGGCGTT 601

QY 819 CCATTGCTTAGGATCGGTCACTGCTGCTCACTCAGTGTTCACCTTCCTGCTCGATCACT 878
DB 602 TGCTGACCCGCGCAGTGATTTCTTGGGCTACGGCGTGTGGGCTTGTGAGTCAAGC 661
QY 879 GGTGATGGGTGTGATCTTGGGTTGTGCGCTTGTGCGCTTGTGCGCTTGTGCGCTTGTG 938
DB 662 CGACGGTGTGCTGTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
QY 939 TTGAGTCGCGAGTTCAAGCGGATTTGGGATCAAGGATATGCGAAGCTTCTTCCAGGCC 998
DB 722 CGGAAGATATGTTAAGCGCGGAGCGCGCATTAAGATAGCGGACATGATTTCCGGAC 781
QY 999 ACGCGGATGTAGACCGTGTGGATGCGATCTCCGCGCGCGATGTGACGTGTTGA 1058
DB 782 ATGCTGCTGATTCGATCGATCGATCGACAGCTGACGCGCGGATACCGCTTTCGCGCTG 841
QY 1059 TCTGAGTGTGATCAGCAGCTCG 1081
DB 842 TGCTACTGTGTCTTTCAGGACG 864

RESULT 8

US-09-221-017B-479/c

; Sequence 479, Application US/09221017B

; Patent No. 6444739

; GENERAL INFORMATION:

; APPLICANT: Ross, Bruce C.

; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF

; NUMBER OF SEQUENCES: 1120

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 755 PAGE MILL ROAD

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows

; SOFTWARE: FastSeq for Windows Version 2.0b

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/221,017B

; FILING DATE: 23-DEC-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PP1182

; FILING DATE: 31-DEC-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PP1546

; FILING DATE: 30-JAN-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PP2911

; FILING DATE: 09-APR-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/AU98/01023

; FILING DATE: 10-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Monroy, Gladys H.

; REGISTRATION NUMBER: 32,430

; REFERENCE/DOCKET NUMBER: 27340-20021.00

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-813-5600

; TELEFAX: 650-494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 479:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7060 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO


```
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...7060
US-09-221-017B-479

Query Match      4.2%; Score 54.4; DB 4; Length 7060;
Best Local Similarity 50.6%; Pred. No. 3.3e-07;
Matches 158; Conservative 0; Mismatches 151; Indels 3; Gaps 1;

QY 725 GATGTGGCGGGTATATCGGGGGTGTCTTTTGGATCGCACCAATGGCGCGTGGTG 784
DB 3245 GATACAGAGCTTTATTCAGGTTCCCTTTTGGCAAGCATACACTCTTCAAGGTCATA 3186
QY 785 AGTCCGAGAGCTTTGGGAAGCTTTCCGGGTCATTTCTTAGATCGGTCACGTGT 844
DB 3185 TCACCCAGAAACATGGGAGGCTTTATCGGAGGCTGTCTCACTGTGGCAGGTGCA 3126
QY 845 GCACTCAGTGTTCCTCTCTCGATCACCACTGTGGTGGTGTGATCTTGGGTGT 904
DB 3125 TTGTGCGAGCCACTATTCGGGTCGGCCAGCATCCGATGGAGCTGGTCTTTTGT 3066
QY 905 GCCTAGTGTGTGC-----GCCAGTTGGGTGACTTGTGTTGAGTCGCAGTTCAAACGGAT 961
DB 3065 CTACTGTGTAACGGCCATGGGCCACTTGGGGCGATCTGTATGAGTCCAACTCTCAAACGCAAT 3006
QY 962 TTGGGCATCAAGATATGTGCAACCTTCTTCCAGGCCAGCGGATTTGATGGACCGTTTG 1021
DB 3005 GCAGGAGTGAAGATTCCGGCCATATCATTCGCCGATCGAGGTATATCTGGACCGGTG 2946
QY 1022 GATGGCATGCTC 1033
DB 2945 GACAGTGTCTCTC 2934

RESULT 9
US-09-540-236-198
; Sequence 198, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 198
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-198

Query Match      4.1%; Score 53.2; DB 4; Length 891;
Best Local Similarity 52.2%; Pred. No. 2.2e-07;
Matches 168; Conservative 0; Mismatches 148; Indels 6; Gaps 2;

QY 723 CGATGTGGCGGGTATATCGGGGGTGTCTTTTGGATCGCACCAATGGCGCGTGG 782
DB 467 CTGATAGTGGGGCTTATTTTATTTGGCAAGAAATTTGGCAACGAAATTTGGCAACCAATG 526
QY 783 TGAGTCCGAGAGCTTTGGGAGGCTT---TGGCGGCTCCATTCCTTAGGATCGGTCA 839
DB 527 TATCACCAATAAAGATATCGAAGGCTTATCGGGTGAATTTGATGGTCTCTGGGGTGGTCA 586
QY 840 CTGGTGCACCTCAGTGTTCACCTTCTCTGCTCGATCACCACTGGTGG---ATGGGTGATCT 896
DB 587 CTGTGGCAGTTGGTTATTTATTTGACGTGTCTGGTATGTCATTTGTTTGAATTT 646
QY 897 TGGGTTGTGCCCTAGTGTGTGGCCACGTTGGGTGACTTGGTTGAGTGCAGTTCAAC 956
```

```
DB 647 TGTCTTGGCTRACTGTACTTGCAGTATACTGGCGCACTTATTTGAATCTATGCTAAAC 706
QY 957 GCATTTGGGCATCAAGATATGTGCAACCTTCTTCCAGGCCACGCGGATTTGAGTCAAC 1016
DB 707 GCCGTGCAGGTATTAAAGATTCTGGGACAATTTCTACCAAGTCAATGGCGGATTTTGGATC 766
QY 1017 GTTTGATGGCATGTCTCCGGC 1038
DB 767 GAATTGATTCGCTACTTTTTCAGC 788

RESULT 10
US-09-596-002-37
; Sequence 37, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 99629
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 78467
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte template ID No. 6632636 37
; PUBLICATION INFORMATION:
US-09-596-002-37

Query Match      4.1%; Score 53.2; DB 4; Length 99629;
Best Local Similarity 52.2%; Pred. No. 4.8e-06;
Matches 168; Conservative 0; Mismatches 148; Indels 6; Gaps 2;

QY 723 CGATGTGGCGGGTATATCGGGGGTGTCTTTTGGATCGCACCAATGGCGCGTGG 782
DB 62838 CTGATAGTGGGGCTTATTTTATTTGGCAAGAAATTTGGCAACGAAATTTGGCAACCAATG 62897
QY 783 TGAGTCCGAGAGCTTTGGGAGGCTT---TGGCGGCTCCATTCCTTAGGATCGGTCA 839
DB 62898 TATCACCAATAAAGATATCGAAGGCTTATCGGGTGGATTTGATGTCTCTGGGGTGGTCA 62957
QY 840 CTGGTGCACCTCAGTGTTCACCTTCTCTGCTCGATCACCACTGGTGG---ATGGGTGATCT 896
DB 62958 CTGTGGCAGTTGGTTATTTATTTTGGAGCTGTCTGGTATGTCATTTGTTTGAATTT 63017
QY 897 TGGGTTGTGCCCTAGTGTGTGGCCACGTTGGGTGACTTGGTTGAGTGCAGTTCAGTCAAC 956
DB 63018 TGTCTTGGCTAACTACTTACTTCCAGTATATCTACCAAGTCAATTTGAATCTATGCTAAAC 63077
QY 957 GCGATTTGGGCATCAAGGATATGTGCAACCTTCTTCCAGGCCACGCGGATTTGATGGTCA 1016
DB 63078 GCCGTGCAGGTATTAAAGATTCTGGGACAATTTCTACCAAGTCAATGGCGGATTTTGGATC 63137
QY 1017 GTTTGATGGCATGTCTCCGGC 1038
DB 63138 GAATTGATTCGCTACTTTTCAGC 63159

RESULT 11
US-09-543-681A-2323
; Sequence 2323, Application US/09543681A
; Patent No. 6605709
```

GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543.681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 2323
LENGTH: 879
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2323

Query Match 3.9%; Score 50.8; DB 4; Length 879;
Best Local Similarity 47.9%; Pred. No. 1.4e-06;
Matches 179; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

QY 668 AATTCATCCCGGTACATATTTTCATTTTGACGTTTCATGCTGTGTGATCGATCGGAT 727
DB 451 AACTCATATACAGGGCTTGTGTGTACTGTATGTCATGTTTGGGCTGGGAC 510
QY 728 GTGGCGGGTATATCGGGGTGTTCTTTGGATCGACCAATGCGCGGTTGGTGGT 787
DB 511 TCAGGGGCTTATGCAATTTGGCCGTTTAAATGGGTAACATAAAATGGCAGAAATATCA 570
QY 788 CCGAAGAGTTTGGGAGGCTTTGCCGGCTCCATGCTTAGTATCGTCACTGGTGCA 847
DB 571 CCGGTAACCTTGGAGGCTGTGGTGGCGGGTAACTACCTGAGGTGTTGTTTCATGG 630
QY 848 CTCAGTGTTCATTCCTGCTCGATCACCACCTGTGTGTGATGCTTGGGTTGTGCC 907
DB 631 CTATTACGGCTTTCGCACCAATTAAGTGCCTAACCATCTATTACTGATTTCCGGC 690
QY 908 CTAGTTGTGCG---CCAGTGGGTGACTTGTGTGAGTCGCGAGTTCAACCGGATTG 964
DB 691 ATCGTGTGATTGTTCCGTATTCGGCGATCTGGCTGAGAGTATGTTTAAACGTGTTCC 750
QY 965 GGCATCAAGGATATGTCAACTTCTTCCAGGCCACGGGATGTAGGACCGTTGGAT 1024
DB 751 GCGATAAAGACAGTAGTCAGCTTATCCCTGGGCGATGGTGTGTAGATCGTATTGAT 810
QY 1025 GGCATGCTCCCGC 1038
DB 811 AGTTTGACCGCGC 824

RESULT 12
US-09-328-352-3131
Sequence 3131, Application US/09/28352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03FA
CURRENT APPLICATION NUMBER: US/09/328.352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 3131
LENGTH: 858
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3131

Query Match 3.5%; Score 45.4; DB 4; Length 858;
Best Local Similarity 56.3%; Pred. No. 9.1e-05;
Matches 85; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 891 TGATCTGGGTGGCCCTAGTCTGTGGCCAGCTTGGGTGACTGGTGGATCGCAGT 950

DB 653 TGATTTTGTCTTTAACTACTACGTGTTGGTTCAGTATTAGGTGATTTATTGAATCAATGA 712
QY 951 TCACACGGATTTGGCATCAAGATATGTGCAACCTTCTTCCAGGCCACGCGGATGA 1010
DB 713 TCACACGGCGGTCTGGTATTAAAGACTCTGCCGCTGTTTCCAGGTCATGGTGGTAT 772
QY 1011 TGGACCGTTTGGATGGCATGCTCCCGGCCG 1041
DB 773 TGGACCGTATTGATCTTTACTCGCAGCAGC 803

RESULT 13
US-09-252-991A-4311/c
Sequence 4311, Application US/09/252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 4311
LENGTH: 1461
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4311

Query Match 3.4%; Score 44; DB 4; Length 1461;
Best Local Similarity 57.1%; Pred. No. 0.00038;
Matches 80; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 909 TAGTTGTGTGGCCACGTTGGTGACTGTTGATCGCAGTTCAAACGCGATTGGCA 968
DB 1405 TGTGTTCTGCTCGATGCTCGCGCACCTTACCGAAAGCATGTTCAAGCCCAATCCGAA 1346
QY 969 TCAAGCATATGTCGAACCTTCTTCAGGCCACGGCGGATTCGACCGTTTGGATGCA 1028
DB 1345 TCAAGCAGCAGCAACCTGCTGCCGCCACGGTGGCGTCTGATCGCATCGACGCC 1286
QY 1029 TGCTCCCGCGCGATGCTG 1048
DB 1285 TGACGGCAGCCATCCCGGTG 1266

RESULT 14
US-09-543-681A-2949
Sequence 2949, Application US/09/543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543.681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 2949
LENGTH: 948
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2949

Query Match 3.4%; Score 43.8; DB 4; Length 948;
Best Local Similarity 46.3%; Pred. No. 0.00034;
Matches 181; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

